PROBLEMS, CURRENT METHODOLOGIES, AND AN ANALYTICAL CASE STUDY IN ELECTRO-ACOUSTIC MUSIC ANALYSIS: LEJAREN HILLER'S *VOCALISE* FROM *SEVEN ELECTRONIC STUDIES FOR TWO-CHANNEL TAPE RECORDER (1963)*

BY

DANIEL SWILLEY

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Musical Arts in Music
with a concentration in Music Composition
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Professor Emeritus Scott A. Wyatt, Chair
Professor Sever Tipei, Director of Research
Professor Jeffrey Magee
Professor Emeritus Chester L. Alwes

Abstract

This dissertation is an investigation of electro-acoustic music analysis; it highlights the difficulties of analysis, explains some of the prevailing analysis methodologies, and demonstrates the use of a hybrid analysis methodology through an analytical case study. These difficulties and analysis methodologies correlate to electro-acoustic music in general as well as the piece chosen for the analytical case study. The piece selected for the study, *Vocalise*, is the first movement from a larger work titled *Seven Electronic Studies for Two-Channel Tape Recorder* (1963) by the composer Lejaren Hiller (1924 - 1994). The study made use of methodologies based on listening (Smalley's Spectromorphology and Roy's Functional and Implicative Method) and used computer-assisted analysis (Music Information Retrieval) as a tool for confirming the findings of the listening analysis.

Acknowledgments

To my mentors Scott Wyatt and Sever Tipei, thank you for your wisdom, patience, and support. I thank my committee members for all their help and feedback. I also thank my friends and family for their support and encouragement.

To my wife Stephanie Koher Swilley, thank you for your unwavering support, patience, exceptional editing skills, and love. Without your inspiration and encouragement, I can truly say I would not be where I am today. And to my two year old son Lincoln, thank you for the much needed distractions, your grins, and your ridiculousness.

Table of Contents

List of Figures

Chapter 1 – Introduction

Electro-acoustic music is a rapidly evolving genre of music with considerable growth over its relatively short history. As a genre, electro-acoustic music has evolved as the technologies that make up the tools of its composition have evolved: magnetic tape recorder, multichannel sound systems, digital audio, computers, etc. In recent years, there has been increasing interest in electro-acoustic music and electro-acoustic music composition, as well as an increase in accessibility to the genre due to the decreasing cost of computer processing power and quality equipment. The analysis of electro-acoustic music has, likewise, increasingly been the topic of research. This research includes scrutinizing the many difficulties associated with the analysis of electro-acoustic music, the formalization and application of analysis methodologies, and contextual aesthetic studies, among others.

The difficulties associated with the analysis of electro-acoustic music are numerous. The lack of a score or other objective representation makes most electro-acoustic music difficult to study in the traditional sense, requiring awareness of the genre's intricacies and alternative methods of analysis. Unfortunately, most music theorists lack the requisite knowledge – specifically the methods, techniques, and technologies of its composition – to engage electro-acoustic music in an analytical context as it differs from acoustic composition.

The analytical writing and examination of contemporary works is a valuable resource to theorists, musicologists, and composers. It fuels the theories of how and why artists engage in their work as well as documents examples of those theories and practices for the next generation of scholars and composers. This document consists of the following: (1) these introductory remarks, (2) a review of the issues concerning the analysis of electro-acoustic music, (3) a review of current electro-acoustic music analysis methodologies, and (4) an analytical case study of the work *Vocalise* (1963) by Lejaren Hiller.

A review of the problems associated with electro-acoustic music analysis and current analytical methodologies is essential to adequately analyze any work of electro-acoustic music. It is important to reassess the state of issues and scholarship often when the subject matter evolves rapidly, as is the case with electro-acoustic music.

Informed by the review sections, the analytical case study is approached as a laboratory for engaging the difficulties and utilizing the methods of electro-acoustic music analysis in the scrutiny of an actual work. The primary focus of the investigation is to identify the structure and primary materials of the work, while negotiating the difficulties of electro-acoustic music analysis. The study made use of methodologies based on listening and also used computer-assisted analysis as a tool for confirming the findings of the listening analysis.

This research topic was chosen for a number of reasons: to better understand the problems and methods of electro-acoustic music analysis, to inform algorithm design in computer-assisted algorithmic composition, and to inform pedagogical approaches in electro-acoustic music composition. The extent of the difficulties and limitations of established methods of electro-acoustic music analysis seemed best examined through an analysis and the appropriate preliminary research. Much of my algorithmic work is based on analysis and the use of a corpus of work or analysis data. The analysis of electro-acoustic music was a step in creating and refining algorithms for sound processing, arranging/mixing, and synthesis. The lexical models offered by many of the electro-acoustic music analysis methodologies based on listening, while designed to address electro-acoustic music in an analytical context, are potential tools for the classroom.

Chapter 2 - Problems in Electro-acoustic Music Analysis

There are many difficulties associated with electro-acoustic music analysis. The problems range from logistical to semantic, and they can all be traced back or linked to one of two primary issues: the absence of a score for most electro-acoustic works and electro-acoustic music's tendency toward musical contexts not centered on the organization of pitches and rhythms. The problems stemming from the lack of a score include a shift in the analysis paradigm from analyzing the score of the work to analyzing the performance, the subjectivity of aural analysis and lack of a prevailing school of thought on aural analysis, and the usefulness of the analysis score and other visualizations (waveform, spectrographs, MIR data, etc.). Issues related to the use of non-traditional musical contexts lead to questions of appropriate descriptive language, what features/materials hold primacy or are important to the work, and the usefulness of traditional analysis models.

2.1 – The Absence of a Score

Perhaps the foremost hurdle in engaging electro-acoustic music in an analytical context is the absence of a score. This lack prompts a re-evaluation of the score as documentation and what information the score provides. The score is an objective representation of a work and the documentation of the composer's intent for a given composition. Concerning the application of traditional scores in electro-acoustic music, Norman Adams states:

> A musical score is a symbolic, or discrete, representation – the score assumes that the music can be abstracted to a sequence of isolated events, or notes. The score is ill suited to visualizing electro-acoustic music because it is often characterized by complex time-varying spectra that defy attempts to be abstracted to discrete events.[1]

The traditional score is not ideal for viewing electro-acoustic music, and yet the score is at the heart of traditional analysis. In acoustic and electro-acoustic music, much can change with each performance of a work: how well the materials were presented in performance, how the materials and instructions in the score might be interpreted in different performances, and how accurately we perceived the nuances of the work. By comparison, the intent present in the score is infallible regardless of what we perceive or fail to perceive in the performance of a work. Without this document, however, our only option is to analyze the performance of a work, or more specifically our perception of the composer's intent based on our aural analysis of the performance.

2.1.1 – Aural Analysis

In the absence of a score, the ear becomes our primary diagnostic tool for the analysis of electro-acoustic music. The approach here, in Camilleri and Smalley's words, is that "the act of concentrated listening, aided by repeated focusing on the same sound, would reveal salient features. As a result, morphologies would be able to be compared and classified, criteria for evaluating individual sounds could be established, and musical

---

[1] N. Adams, "Visualization of Musical Signals," in *Analytical Methods of Electro-*

‘values’ could be ‘abstracted’.”[2] This has been the core analytical strategy of almost all electro-acoustic music analysis systems that have appeared since Pierre Schaeffer’s *Traité des objets musicaux* (1966).[3]

Many useful observations can be made about a work through aural analysis, however, the inherent subjectivity of an examination that is based solely on listening must be scrutinized. As John Young asserts, an “aural analysis of music in its *concrete* form invites the criticism that it is too filtered by what we think we want to hear…”[4] The training and listening capabilities of the individual listener/analyst can be quite varied. The analyst here becomes an undefined and unknowable filter on the analysis of a work – someone reading such an analysis cannot know to what level the analyst accurately perceives the nuances of a work. This uncertainty in the quality of the investigation is an unavoidable fact of electro-acoustic music analysis. At present, it can only be combated by the acknowledgement that the given examination is simply one perspective on the work and by, if possible, seeking out multiple analyses of the work and forming a composite of the various analyses to understand the intricacies of the work.

Furthering the difficulties of aural analysis, there is no traditional methodology or model to use as a framework for the aural analysis of electro-acoustic music. Most musicians are trained to identify and dictate intervals, chords, melodies, harmonies, and

---

[2] Camilleri, L., and D. Smalley. “The Analysis of Electro-acoustic Music: Introduction.” *Journal of New Music Research* 27.1-2 (1998): 4.
[3] Schaeffer’s treatise focuses on sound typologies, and provides a theoretical framework for identifying and classifying the morphological features of sound. The first such document that identifies issues and strategies for electro-acoustic music analysis.
[4] Young, J. “Sound Morphology and the Articulation of Structure in Electro-Acoustic Music.” *Organised Sound* 9.1 (2004): 8.

rhythms. However, the usual extent to which musicians are trained in the aural analysis of music is the identification of phrase structure and cadences in the study of small forms (simple binary, rounded-binary, and ternary). Composers often engage in more rigorous analytical listening in the context of literature/score study, but are rarely doing a formal analysis in acoustic or electro-acoustic contexts.

2.1.2 – The Analysis Score

An analysis score is a tool often utilized in the examination of electro-acoustic music. Lacking a traditional score on which to annotate with the results of the analysis, as in traditional analysis, the analysis score is used as a score substitute that aids in documenting one's scrutiny of a work. The analysis score becomes a visual representation of the work – a graphism that typically takes the form of an elaborate timeline where primary sonic features are represented with appropriately shaped notations. The information conveyed in an analysis score can vary greatly, depending primarily on what information is central to the work, but also on the particular focus of the examination being performed. Some analysis scores are drawn freehand, yet many take the form of an annotated sonogram[5] as the basis for illustrating the desired features of the work.

It is common to find formal or proportional diagrams in traditional analysis, but the analysis score is typically not a tool brought to bear in the examination of works that

---

[5] A computer visualization to be explained further in section 2.1.2.2.1 – Common Visualizations

have a score. There are reasons for the analyst to sometimes make a reduction of a work for the sake of simplification, or isolate the important materials or other information. The foremost example of this is the *Schenkerian* graphism used to illustrate the tonal function and structure in tonal music. The merit of such a hierarchical graphism is clear, however, similar graphisms have not made their way into the analysis of electro-acoustic music where the evolution of pitch was not significant.

2.1.2.1 – A Score: *A Priori* vs. *A Posteriori*

The analysis score and other non-composer derived visualizations, unfortunately, suffer from the same issues of subjectivity previously discussed in reference to aural analysis. A score not penned by the composer, is a score that bears the interpretation of another and may not accurately relay the composer's intent for the work. The composer's score is *a priori* to the performance of the work, while non-composer derived visualizations are *a posteriori* to the performance. Again, the inherent subjectivity involved here is not grounds for disregarding all such representations, they must simply be acknowledged as a perspective.

It is also important to note that any visual representations of music are invariably flawed. Music is a time-based art form experienced aurally; no notational scheme or method of visualization has yet been invented that adequately express all nuances of music. Most contemporary composers lament the limits of Common Practice Notation. The search for a perfect visualization that adequately represents all that is expressed in music is, perhaps, an unrealistic one. Musical representations must be scrutinized for what information is not conveyed as well as for what information is conveyed. As Mary Simoni notes, "musical abstractions, once harnessed by some representation, are invariably cheated of some aspect of their meaning."[6]

2.1.2.2 – Other Visualizations

The ear is our most important tool for the scrutiny of works lacking a score, but there are a number of useful computer tools that can aid the analysis. There are many different types of analysis algorithms that can be implemented allowing the extraction of data and generating visualizations that could illustrate myriad aspects of a work. As John Young states, "analytical software tools […] enable a kind of extension of the ear – to isolate and magnify elements of a spectrum, for instance."[7] Many useful observations can be made from the simplest and most common visualizations (waveform and spectrogram,) and yet greater insights are possible via Music Information Retrieval (MIR) analysis (discussed in section 2.12.2.2). Such visualizations are often utilized in constructing an analysis score.

2.1.2.2.1 – Common Visualizations

As Norman Adams asserts, "a visualization method that translates the aurally salient qualities of electro-acoustic music into a static image is important to an

---

[6] M. H. Simoni, "Introduction," in *Analytical Methods of Electro-acoustic Music*, ed. M. H. Simoni (2006): 7.
[7] Young, J. "Sound Morphology and the Articulation of Structure in Electro-Acoustic Music." *Organised Sound* 9.1 (2004): 7.

understanding of electro-acoustic music."[8] The most common visualizations for audio are the Time-Domain Representation, Frequency-Domain Representation, and the Time-Frequency-Domain Representation (or Sonogram.) As expected, each of these methods of visualizing audio signals has advantages and limitations.

The Time-Domain Representation or Waveform Display is perhaps the most common visualization used to display audio data. As displayed below in Figure 1, it is a two-dimensional display of the fluctuations of the air pressure wave over time, with time as the horizontal axis (x) and amplitude as the vertical axis (y).[9] It is the common visualization used in Digital Audio Workstations (DAWs) as well as Audio

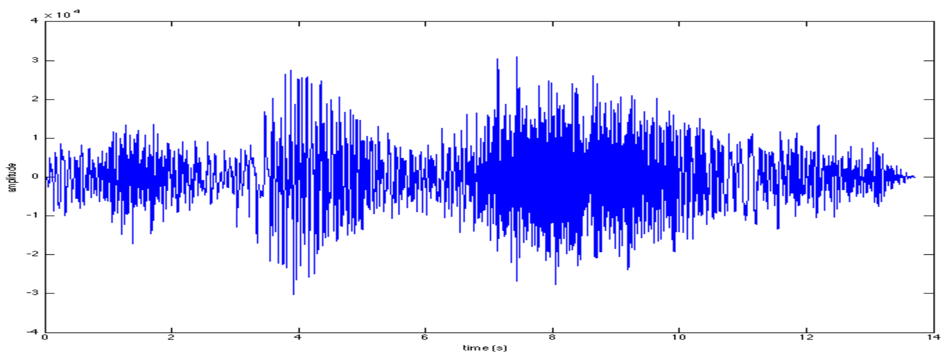


Figure 1 – Time-Domain Representation or Waveform Display example.

Editor applications, and is useful for gaining a sense of the acoustic energy present in an audio file. Details of the Time-Domain Representation translate the activity of the audio signal, yet there are some instances when the visualization "does not correlate well to

---

[8] N. Adams, "Visualization of Musical Signals," in *Analytical Methods of Electro-acoustic Music*, ed. M. H. Simoni (2006): 13.
[9] Ibid.. 14.

aural perceptions of the waveform."[10] In addition, there are instances where the Waveform Display does not successfully represent all layers of material present in an audio signal – the masking of materials, even prominent materials, can occur quite easily. The previous Figure 1 exemplifies this phenomenon. To illustrate this, Figure 2 is a display of the same audio file from Figure 1 filtered (or decomposed) into ten separate audio streams of different frequency registers from low to high. The filtering reveals much activity in the upper registers that is not apparent from an examination of the Figure 1 waveform alone.

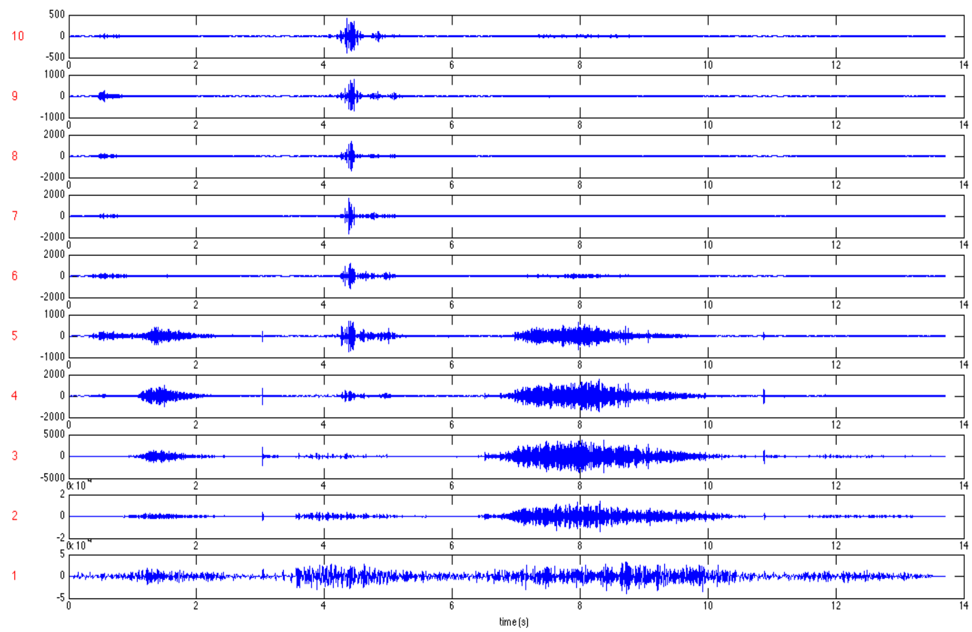


Figure 2 – Masking revealed by filtering

The Frequency-Domain Representation or Spectrum Display is a common visualization utilized when analyzing audio for its frequency content. It is a two-

---

[10] Ibid.. 15.

dimensional display of the relative strength of frequencies present in an audio signal with (typically) frequency as the horizontal (x) axis and relative strength as the vertical (y) axis. The Spectrum Display is typically used to display the frequency content of a specific moment of an audio file, as in Figure 3, or to run in real-time displaying the frequency content of an input signal. It can also be used to display the average frequency content for a defined window of time (e.g. a phrase, the first 30 seconds, or the whole audio file as in Figure 3 – [Right]). The strength of the Frequency-Domain Representation is also its weakness. It is useful to be able to scrutinize the frequency content at a given moment or over the span of any part of an audio file; however, music

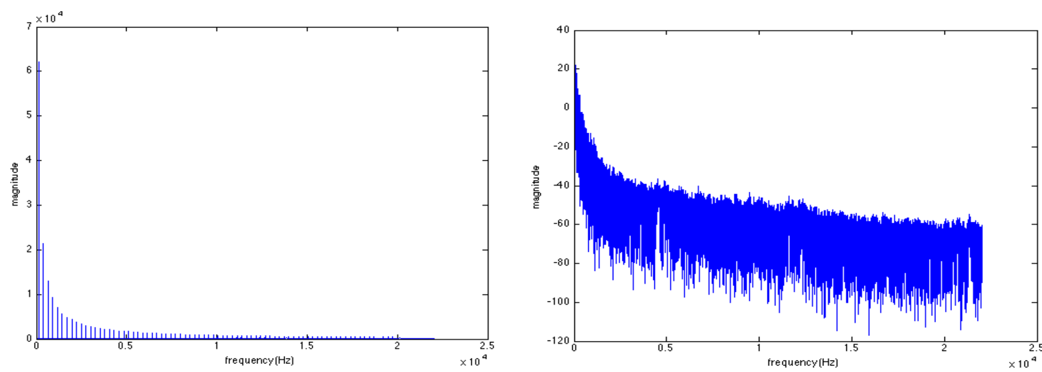


Figure 3 - [Left] Square wave and [Right] Figure 1 Audio file average Frequency-Domain Representations

is a time-based medium experienced as a continuum, and a visualization that does not express this continuum of time fails in its representation of music. The Frequency-Domain Representation remains a useful tool, akin to a magnifying glass, for the scrutiny of frequency content.

The Time-Frequency-Domain Representation or Sonogram is a very useful visualization that can be generated by modeling the audio signal as a collection of sinusoids.[11] The Sonogram bears the strengths and none of the weaknesses of both the Time-Domain and Frequency-Domain Representations: the times and frequencies at which a signal has energy can be identified with detail, without the issues of masking[12] and loss of the continuum of time as in Time-Domain and Frequency-Domain

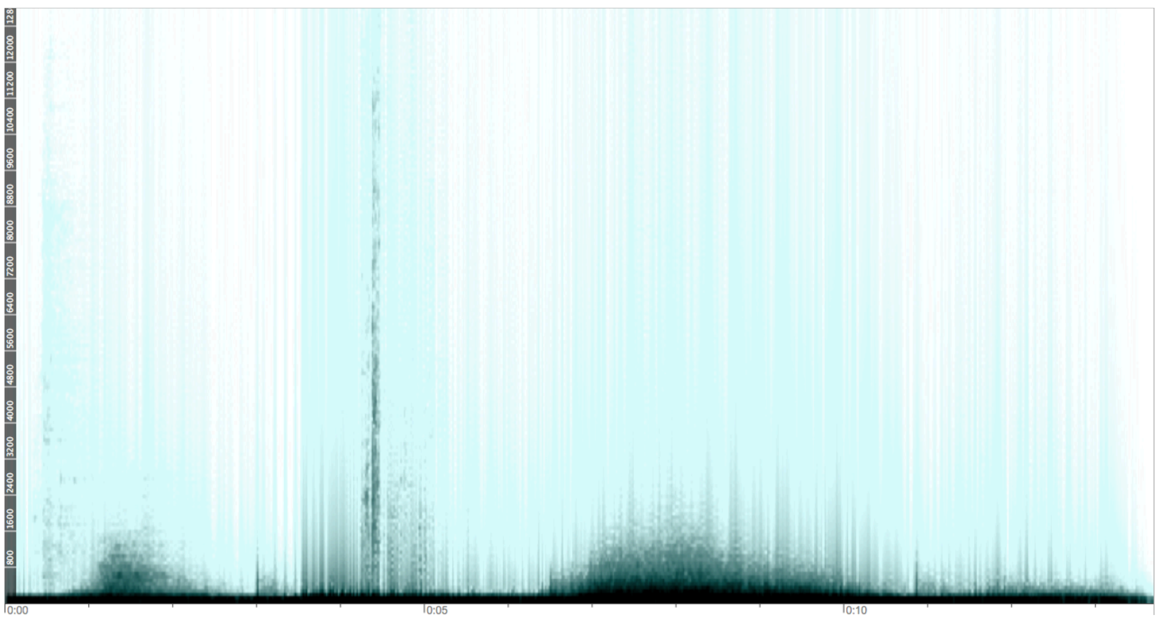


Figure 4 - The Sonogram

Representations respectively. As in Figure 4, the Time-Frequency-Domain Representation is typically a two-dimensional visualization with time as the

---

[11] N. Adams, "Visualization of Musical Signals," in *Analytical Methods of Electro-acoustic Music*, ed. M. H. Simoni (2006): 14.
[12] For clarification, without masking by different frequencies present in a signal thus causing a lack of correlation between the visualization and the oral experience, as in Time-Domain Representations.

horizontal axis, frequency as the vertical axis, and a third dimension indicating the amplitude or intensity of that particular frequency/time by means of a colormap. Colormaps are inherently poor at displaying accuracy in visualizations. Most software allow a fair amount of flexibility in adjusting the colormap so that it best suits the audio to be displayed: the dynamic range to be displayed, color scheme (a gradient of colors or multi-hued scheme), and options for how to display multiple channels of audio. Figure 5 shows some of the variation in colormaps.

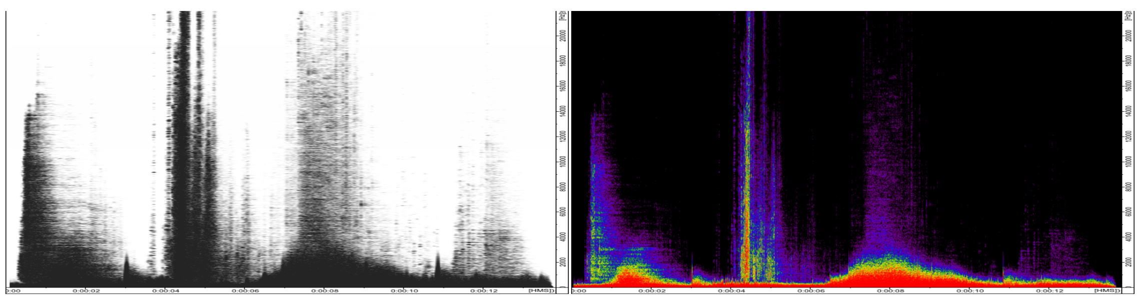


Figure 5 – [Left] Poorly adjusted gray-scale and [Right] Classic Rainbow Sonograms

The Sonogram is not without some other limitations such as the matter of adjusting the parameters of the analysis used to generate the visualization so that it best displays the desired features of the audio. This can be challenging due to what is called the uncertainty principle. "The uncertainty principle states that as the time resolution decreases, the frequency resolution increases and vice versa."[13] This is an unavoidable aspect of utilizing the Short Time Fourier Transform, which is the type of analysis

---

[13] N. Adams, "Visualization of Musical Signals," in *Analytical Methods of Electro-acoustic Music*, ed. M. H. Simoni (2006): 20.

utilized in creating a Sonogram. The issues of creating visualizations based on spectral estimation are illustrated in Figure 6. Despite the limitations and need for adjustments to best display the content of an audio signal, the Time-Frequency-Domain Representation is among the best tools for the analysis of audio content or electro-acoustic music.

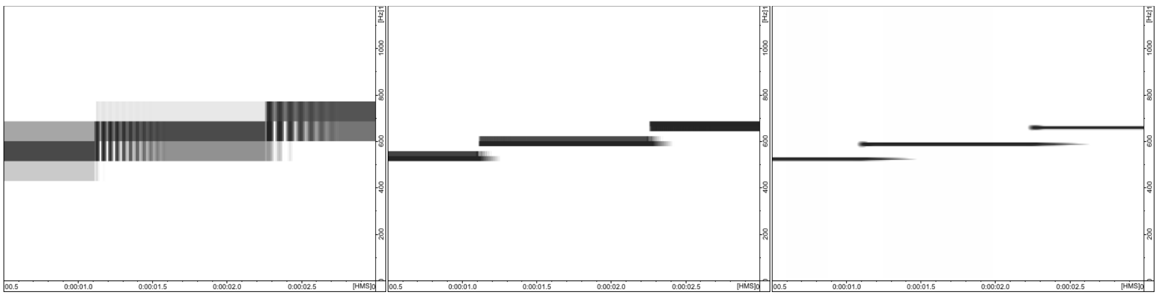


Figure 6 – Three Sonograms of the same audio with different analysis setting: [Left] displaying the best time and worst frequency resolution; [Right] displaying the best frequency and worst time resolution; and [Center] displaying acceptable resolution for both frequency and time.

2.1.2.2.2 – MIR Data Visualizations

Although less common than the Waveform or Sonogram displays, there is an increasing number of software tools that allow the analysis and subsequent generation of visualizations of audio signals via methods of Music Information Retrieval. There are a few exceptions, but most MIR analysis and feature extraction is done via programing or scripting of some kind. This coding allows great flexibility in how the analysis will be directed and how the extracted data will be displayed. In this light, visualizations generated via MIR can vary considerably: two-dimensional plots of an extracted feature's evolution (y) over time (x); three-dimensional plots of two extracted features' evolution over time with one of the extracted features displayed on the z axis as in Sonograms; and three-dimensional plots of two or more extracted features. MIR visualizations have great potential for displaying the evolution of significant aspects of a given musical example. The use of MIR feature extraction in the analysis of electro-acoustic music – including software options and extractable features – will be addressed further in the following chapters.

2.2 – Non-traditional Musical Contexts

Electro-acoustic music's tendency toward non-traditional musical contexts is a somewhat controversial issue. As John Young states, "a view of musical sounds as complex morphological events and not simply 'notes' of specific pitch center and duration is an innovative musical observation…"[14] The emergence of electro-acoustic music has broadened the definition of music; the definition of music analysis should be likewise adjusted. The evolution of music analysis must include new terminology and descriptive language to appropriately address the features and aspects of electro-acoustic music. Similarly, analysts must familiarize themselves to some degree with the various aesthetics of electro-acoustic music and learn how to identify the salient aspects of a work. A considerable knowledge base is necessary to adequately engage the non-traditional aspects of electro-acoustic music analysis, however, such analysis does not require the abandonment of traditional analysis models (many of which could be utilized in the scrutiny of a musical discourse no matter the medium, genre, or aesthetic).

---

[14] Young, J. "Sound Morphology and the Articulation of Structure in Electro-Acoustic Music." *Organised Sound* 9.1 (2004): 9.

2.2.1 – Language

The use of appropriate descriptive language in the analysis of electro-acoustic music is a serious issue; an analyst must be able to appropriately describe the subject matter of their analysis. There are many traditional music terms that can be used to describe aspects of an electro-acoustic composition, yet there are some that simply do not translate. Several composers of electro-acoustic music have contributed lexicons of descriptive terminology for electro-acoustic music and its features: Pierre Schaeffer's Typomorphology, Denis Smalley's Spectromorphology, and Stéphane Roy's Functional and Implicative Analysis. These lexical models are useful, but there are inevitably unforeseen situations that are best represented by the logical extension of existing models and descriptive terminologies.

2.2.2 – Salient Features and Materials

The determination of salient materials/features of an electro-acoustic work can differ from that of an acoustic work in the type and extent that musical parameters are involved: meter, pitch, and timbre. The parametric approach of traditional analysis, unfortunately, does not always fit well with the scrutiny of electro-acoustic music. The salient materials/features of electro-acoustic music can include any number of possibilities (e.g. the evolution of a specific gesture, source material, or textural idea, as well as the interaction of processed and unprocessed source materials, synthesized and recorded source materials, or different spatialization trajectories). It is sometimes more accurate to think of a work as having a central idea or trajectory rather than specific materials that are being developed. As Cammilleri and Smalley state, "An important goal of analytical exploration is… to attempt to reconcile and relate the internal world of the work with the outside world of sonic and non-sonic experience."[15] It is arguably an important aspect of scrutinizing some works, however, this does not mean that interpreting an electro-acoustic work is primarily an exercise in identifying source materials. As Leigh Landy asserts, "Source recognition forms only part of the understanding of a work, and in fact, may impede understanding. Nevertheless, real-world sound references also form part of the communality of experience…"[16]

2.2.3 – Traditional Analysis

Many aspects of traditional analysis models fail to adequately address the materials under examination in electro-acoustic music, yet there are many extrapolations and extensions of those models that are useful. The concepts in traditional analysis that contribute little in electro-acoustic music analysis mostly pertain to the scrutiny of pitch and harmony, which is the focus of an average music student's studies in the examination of western music. Many of the theories related to form, function, and the transformation of materials can be utilized via extrapolation taking into consideration the transformative processes used in the construction of electro-acoustic music materials. There are some

---

[15] Camilleri, L., and D. Smalley. "The Analysis of Electro-acoustic Music: Introduction." *Journal of New Music Research* 27.1-2 (1998): 5.
[16] L. Landy, "The Intention/Reception Project," in *Analytical Methods of Electro-acoustic Music*, ed. M. H. Simoni (2006): 30-31.

types of materials found in acoustic music that traditional analysis models have difficulty with: timbral evolution, textural materials, and gestures. This is unfortunate, as these are primary areas of interest for most electro-acoustic music composers due to the practically unlimited possibilities of transformation and layering available in electro-acoustic music. Eventually, electro-acoustic music analysis will not be viewed much differently from the analysis of acoustic music in terms of the difficulties associated with addressing musical contexts not centered on the organization of pitches and rhythms. As Rosemary Mountain states:

> There are many techniques and musical structures which are often seen as typifying electro-acoustics, but are found increasingly in acoustic works as well: juxtapositioning and layering; spatialisation; timbral modulation; microtones, continua, and other alternatives to equal-tempered organizations of frequencies; development of musical textures; and gesture.[17]

---

[17] Mountain, R. "Theories Market: Open for Trading." *Organised Sound* 9.01 (2004): 17.

Chapter 3 - Current Methodologies in Electro-acoustic Music Analysis

Electro-acoustic music analysis, as yet, has no codified, universally applicable analysis methodology. Electro-acoustic music is a broad genre consisting of many varied subgenres including tape music, computer music, soundscape, sound art, interactive electronics, electronic music with acoustic instruments, and laptop music. The development of a universally applicable analysis methodology is likely never to come about and would, perhaps, be an undesirable goal. Should such an all-encompassing methodology develop, it would undoubtedly fail to yield a suitable level of detail and depth of analysis in all types of electro-acoustic music.

Although we lack a standardized universally applicable methodology, there are a number of methodologies currently employed in electro-acoustic music analysis. Most methodologies are based on listening due to the difficulties presented by the lack of score. There is, however, much research being done in computer-assisted analysis utilizing Music Information Retrieval feature extraction techniques, which avoid the subjectivity of a listening based examination. Music Information Retrieval instead suffers from issues of subjectivity in the interpretation and guiding of the analysis. Among the methodologies based on listening, there are a variety of approaches based on identifying different aspects that are significant to the given work: Denis Smalley's Spectromorphology, focused on describing the spectral content[18] of sound objects[19];

---

[18] Spectral content refers to the particular component frequencies and amplitudes that make up a sound. The spectral content of a sound determines its timbre.
[19] A sound object is a sonic event used in the composition of an electro-acoustic work, equivalent to a note in terms of scope.

Stéphane Roy's Functional and Implicative Analysis, focused on the interrelationship of sound objects; and Robert Frank's Temporal Elements, focused on the perceived stability of musical contexts based on the repetition and rhythm of materials.

3.1 - Denis Smalley – Spectromorphology

Denis Smalley's Spectromorphology is a long-standing and well-known analysis methodology and descriptive tool for electro-acoustic music. The term *Spectromorphology* is a fusion of the words spectra and morphology. This methodology is a system of analysis concerned with the spectral content of sound objects and how spectral content changes over time. In Smalley's words, Spectromorphology is, "an approach to sound materials and musical structures which concentrates on the spectrum of available pitches and their shaping in time."[20] In 1997, Smalley states this analysis framework is to be used for, "describing and analyzing the listening experience."[21] Additionally, Smalley asserts that Spectromorphology is:

> …intended to account for types of electro-acoustic music which are more concerned with spectral qualities than actual notes, more concerned with varieties of motion and flexible fluctuations in time rather than metrical time, more concerned to account for sounds whose sources and causes are relatively mysterious or ambiguous rather than blatantly obvious.[22]

---

[20] Smalley, D. "Spectro-Morphology and Structuring Processes." *The language of electro-acoustic music* (1986): 61-93.
[21] Smalley, D. "Spectromorphology: Explaining Sound-Shapes." *Organised Sound* (1997): 107.
[22] Ibid., 109.

Smalley's array of descriptive tools for classifying the spectral content of a sound object is organized into the following: the Note to Noise Continuum, the Occupancy of Spectral Space, and the Spectral Density. The Note to Noise Continuum is the general descriptor range for the spectral content of a sound object. This approach is optimal as many of the events/materials in electro-acoustic music are difficult to define in a finite way due to their continuously evolving makeup. Many of Smalley's other descriptors are similarly approached as a continuum of possibilities between two opposing descriptors (antonyms). Further, general modifiers are employed to describe the qualities of a sound object: the *harmonicity* to *inharmonicity continuum*, and classification of *granular* or *saturate* Noise events. The Occupancy of Spectral Space refers to how a sound object may be perceived, "over the spectrum of audible frequencies."[23] Smalley defines three types of spectral space: *canopy*, *centre*, and *root*. The canopy and root act, as the names suggest, as upper and lower limits that frame the centre. There are four spectral space qualifying descriptors, using antonym pairs: *emptiness*/*plentitude*, *diffuseness*/*concentration*, *streams*/*interstices*, and *overlap*/*crossover*. Spectral Density encompasses the *distant* to *close continuum* as well as masking of sound objects. Smalley has defined six spectral density types: *filled*, *packed*, *compressed*, *opaque*, *translucent*, *transparent*, and *empty*.

Smalley's framework of descriptive tools for classifying the morphology of a sound object is organized into the following categories: (1) Onsets, Continuants and Terminations functions; (2) Motion and Growth; (3) Texture Motion; and (4) Behaviour.

---

[23] Ibid., 118.

The Onset, Continuant, and Termination functions use the descriptors for the evolution of events, gestures, and textures on the micro and macro scale listed in Figure 7. Smalley

Onsets	**Continuants**	**Terminations**
Departure	Passage	Arrival
Emergence	Transition	Disappearance
Anacrusis	Prolongation	Closure
Attack	Maintenance	Release
Upbeat	Statement	Resolution
Downbeat		Plane

Figure 7 – Terminology of Onsets, Continuants, and Terminations.

defines Motion and Growth processes with categories of *unidirectional*, *reciprocal*, *cyclic/centric*, and *bi/multidirectional* – each category having its own body of descriptive terminology. In regard to Texture Motion, Smalley states, "Most of the bi/multidirectional motions imply internal textural change,"[24] and "in terms of occupancy of spectral space they could vary in dimensions, and consist of more than one layer."[25] Smalley defines "ways in which the internal textural components may collaborate in motion,"[26] along with sufficient modifying terminology for a variety of textural motions: *streaming*, *flocking*, *convolution*, and *turbulence*. Lastly, behavior is used to describe the relationships between Spectromorphological entities in a given musical context. This is similar to the relationships a seasoned listener could identify in music for acoustic instruments: *dominance/subordination*, *conflict/ coexistence*, *equality/ inequality*,

---

[24] Ibid., 117.
[25] Ibid.
[26] Ibid.

reaction/interaction/reciprocity, *activity/passivity, activity/inactivity*, and *stability/instability*. In addition, Smalley uses the *loose-tight continuum* to describe the "degree of coordination freedom,"[27] and the *voluntary-pressured continuum* to classify relationships of how materials develop over time.

Addressing the classification of space and localization in electro-acoustic music, Smalley uses "the term *spatiomorphology* to highlight the special concentration on the exploration of spatial properties and spatial change, such that they constitute a different, even separate category of sonic experience."[28] Spatiomorphology is organized into two categories: *listening space* and *composed space*. The *listening space*, having to do with a listener's position and atmosphere relative to the "frontal image,"[29] is divided into *diffused* and *personal space.* This includes the modifiers: *intimacy/distancing*, *breadth/depth*, *image definition/localization*, *orientation*, and *spectral quality*. The *composed space* is likewise divided into two categories: *internal* and *external space.* Smalley states, "*internal space* occurs when a Spectromorphology itself seems to enclose a space,"[30] and uses the examples "hollow wooden resonance, metallic resonance, stringed instrument pizzicato resonance,"[31] citing that their sonic qualities give the impression that their enclosed within more solid materials. *External space*, the more significant aspect of *spatiomorphology*, has an elaborate list of sub-categories and descriptive terminology for the many properties of spatialized sound: *image definition*

---

[27] Ibid., 118.
[28] Ibid., 122.
[29] Ibid.
[30] Ibid.
[31] Ibid.

(*concentrated-diffuse*, *blurred-clear*) and *spatial fill* (*plentitude-emptiness*) under the sub-category of *focus*; *distribution style* (*isolation*, *exchange patterns-groupings*, *scattering*) under the sub-categories of *non-contiguous space* and *spatial texture*; *spread settings* and *trajectories* (*paths*, *velocities*, *residues*) under the sub-categories of *contiguous space* and *spatial texture*; and characteristic paths such as *approach*, *departure*, *crossing*, *rotation*, and *wandering*.

Smalley's Spectromorphology offers much for the electro-acoustic music analyst in terms of lexical fodder as well as descriptive and organizational concepts. The terms vary in usefulness, however, the concepts of a continuum between two descriptors and how to divide the aspects of a sonic event are universally applicable in electro-acoustic music and quite useful. Smalley's methodology serves as a good model for descriptive language, although it only minimally addresses the relationships of materials by his inclusion of the behavior term pairs.

3.2 - Stéphane Roy – Functional and Implicative Analysis

Drawing on the writings of Leonard B. Meyer, Stéphane Roy's Functional and Implicative Analysis focuses on describing the contextual interrelationships of sound objects. In Roy's words, his methodology is "analytical and interpretative,"[32] and "inspired by the functionalism of language, that is, the fact that the role of one semantic

---

[32] Roy, S. "Functional and Implicative Analysis of Ombres Blanches*." *Journal of New Music Research* (1998): 166.

unit can change according to its location in the syntactic flow."[33] The analysis begins with the segmentation of a work, based on aural perception, into what Roy refers to as "syntactical units."[34] The functional analysis is then addressed via a defined lexicon of functions, which are "ascribed to the units according to their perceived contextual roles."[35] Roy adapts Meyer's implicative method for tonal melody analysis to address implicative qualities of units in electro-acoustic music. Roy uses Meyer's definition of implication, which "is an hypothesis that a competent listener makes about the progress or possible resolution of a pattern, based on inferences deduced from the context."[36] More over, in Meyer's words:

> An implicative relationship is one in which an event – be it a motive, a phrase, and so on – is patterned in such a way that reasonable inferences can be made both about its connections with preceding events and about how the event itself might be continued and perhaps reach closure and stability.[37]

To aid in the analysis, Roy makes use of a listening/analysis score – a graphic representation of the syntactical units of a work identified during the segmentation process. Roy defines a set of symbols for representing both the functional and implicative relationships of a work in the analysis/listening score. Abbreviations of the function names would have adequately identified the function of analyzed units in the listening/analysis score, yet Roy's symbol set includes graphics for each function. The

---

[33] Ibid., 166.
[34] Ibid., 165.
[35] Ibid.
[36] Ibid., 180.
[37] Meyer, L. B. *Explaining Music*. Univ of California Press, 1973. 110.

function symbol would be placed next to the graphic representing a given unit of a work, making use of arrows if space is limited in the score. The graphism used to label implicative relationships, Figure 8, takes the form of a bracket of sorts with either end being placed on the related units. Features of the graphic are altered to show different implicative relationships: arrows to the right for an explicit event and arrows to the left for the inverse implications.

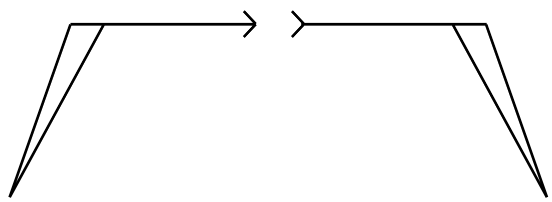


Figure 8 - Roy's Implicative Graphic.

Roy's functional analysis, which contains a total of forty functions, is organized into five categories: orientation, stratification, process, rhetoric, and rhythm. Below is a partial list of Roy's functions with symbols and definitions:

Orientation Category:

> <u>Begetting</u> (or) – a dynamic function characterized by a brief morphological gesture that prepares and furthers following event(s).[38]
>
> <u>Conclusion</u> ( ) – a consequent function that terminates a phrase or section without ambiguity – requires preparation.

---

[38] Roy, S. "Functional and Implicative Analysis of Ombres Blanches*." *Journal of New Music Research* (1998): 181.

Interruption (⇥) – a function that is unprepared and unresolved – "morphological rupture."[39]

Introduction ( ) – a function that "progressively initiates a musical discourse or part of it by using a dynamic crescendo, an increase in density, or any kind of gradual morphological progression that does not generate surprise in the listening process."[40]

Suspension ( ) – a weak consequent function; similar to the *Conclusion*, but lacking in stability.

Trigger (↦) – an unprepared function that, like the *Interruption*, breaks the morphological flow; causally linked to a consequent.

Stratification Category:

Background ( ) – the lowest hierarchical level function, contributing only low-level information characterized mostly by stasis – too long to be memorized.

Figure (▲) – a function characterized by being short and well-articulated at the *Foreground* level.

Foreground ( ) – the highest hierarchical level function, characterized by "well-articulated sound unit with a duration much longer than that of the *Figure*"[41] - too much high level information to be memorized.

Rhetorical Category:

Affirmation (>!) – a function that terminates a process of repetition – related to *Conclusion* for its terminating characteristic and *Reiteration* for its antecedent/consequent role.

Announcement (>A>) and Reminder (<R<) – "these functions are represented by a very prominent sound unit. Usually, an *Announcement* states a fragment of

[39] Ibid., 181.
[40] Ibid.
[41] Ibid., 182.

the *Reminder* in order to prepare for, and to increase the perpetual importance of the complete event that will appear later."[42]

Call (?C>) and Answer (>A!) – these functions are "based on a local rhetorical relationship, and [are] articulated through the repetition of an expressive antecedent/consequent pair. When a few repetitions happen within a short time, the *Call* function becomes associated, by conditioning, with the *Answer."*[43]

Deflection (> ∨) – a function which interrupts the continuity of the primary morphology to pursue a new goal.

Parenthesis (( )) – a function that "is represented by an encrustation, that is to say by a sound unit or a group of sound units that temporarily break into a musical progression without having any causal motive."[44]

Reiteration (>>>) – a function characterized by the use of frequent repetition of events.

Sign (< ›) – the function for sound units that trigger extra-musical references.

Rhythmic Category:

Pedal (⎍⎍⎍) – a function associated with long sound units that alter the perception of tempo.

Stéphane Roy's methodology addresses the need for terminology to scrutinize the relationships of sound events in electro-acoustic music. Additionally, it is beneficial to see an example of how Meyer's implicative theories for melody are modified for use in electro-acoustic music. A model of addressing functional and implicative relationships is extremely useful.

---

[42] Ibid., 182.
[43] Ibid., 183.
[44] Ibid., 183.

3.3 - Robert Frank – Temporal Elements

Frank's Temporal Elements is a cognitive-based system "meant to be a first-level tool for theorists, composers, and teachers […] applicable to music regardless of the presence or absence of pitched-based materials, meter, or notation."[45] It is a system derived from how our minds perceive music in what W. Jay Dowling calls the "psychological present" – our internal negotiation of musical materials, where new or unexpected materials are foremost, and regular or expected materials are background. In Frank's words, the "interplay of consistency verses [*sic*] change results in the perception of, and transformations between, foreground and background material and is crucial to understanding the pacing and temporal unfolding of events in a composition."[46] Frank's system highlights what we hear as stability and instability in music with the following classifications or Temporal Elements: Sustaining, Aligned/Repeating, Aligned/Non-Repeating, Non-Aligned/Repeating, and Non-Aligned/Non-Repeating. These classifications illustrate Frank's differing views concerning repetition as compared to Dowling – who viewed all repetitive patterns equal to sustained sounds to the listener.

Frank's five basic Temporal Elements for classifying sounds and sound masses are defined as follows:

> <u>Sustaining (S)</u> – consisting of sustained sound or held tone of non-inconsiderable duration (e.g. – a held fermata, pedal point, or lengthy resonance).

---

[45] Franks, R. "Temporal Elements: a Cognitive System of Analysis for Electro-Acoustic Music." *ICMC 2000* (2000): 1.
[46] Ibid., 2.

Aligned/Repeating (A/R) – repeated material coordinated with a regular pulse (e.g. – common accompanimental materials or materials associated with a minimal aesthetic).

Aligned/Non-Repeating (A/NR) – non-repetitive material coordinated with a regular pulse. Can consist of "unison passages, homogenous chorales, and metered music with distinctly independent lines with metrically congruent subdivisions of the beat."[47]

Non-Aligned/Repeating (NA/R) – repetitive materials lacking coordination to a regular pulse (e.g. – textural materials containing independently evolving repetitive events).

Non-Aligned/Non-Repeating (NA/NR) – non-repetitive material lacking coordination to a regular pulse. Can consist of a wide variety of materials in contemporary musics.

Figure 9 is Frank's Temporal Elements table, which illustrates the continuum in the five basic Temporal Elements. The table is read, from most stable to least stable: Sustains (S), Aligned/Repeating (A/R), Non-Aligned/Repeating (NA/R), Aligned/Non-Repeating (A/NR), and Non-Aligned/ Non-Repeating (NA/NR).

[47] Ibid., 3.

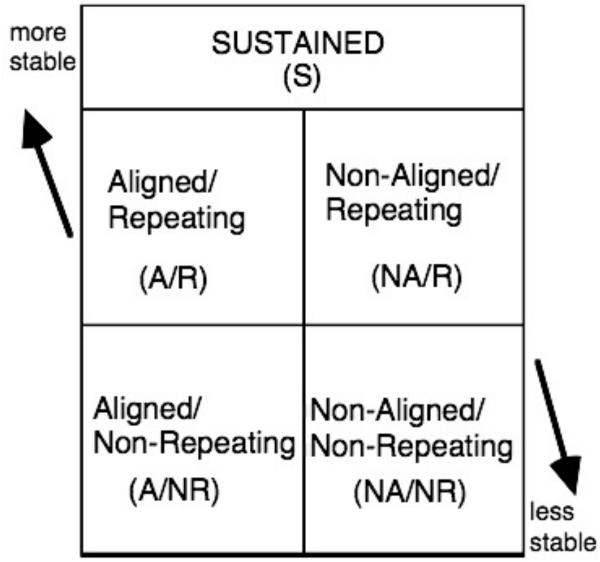


Figure 9 - Frank's Temporal Elements Table.

Frank addresses complexity in the temporal evolution of a work in terms of the juxtaposition of Temporal Elements and transformation of Temporal Elements. Temporal Elements can exist singularly or in juxtaposition, combining similar or different Temporal Elements in various presentations. The two most common situations: (1) the simultaneous starting of two or more Temporal Elements, creating a period of disorientation as the listener determines which materials to focus on; and (2) the gradual layering of Temporal Elements of similar or disparate relations for mass effect. Frank defines two categories of transformations between Temporal Elements: (1) Hybrid Elements, "combining traits of two different types of Temporal Elements;"[48] and (2) Gradual Transformations, consisting of a gradual change of the materials to the point of needing to adjust the classification.

[48] Ibid., 4.

Frank presents his system as a tool for understanding the unfolding of temporal aspects of a work, but states that the categorizations he has defined are, “intended to be general, not absolute or all-inclusive.” [49] There are some ambiguities to negotiate, such as the limits of the classifications and areas of temporal flux/ambiguity. An analyst must bear these in mind and evaluate on a case-by-case basis what insights may be gleaned by the use of this system.

3.4 - SQEMA and Music Information Retrieval

The SQEMA, or Systematic and Quantitative Electro-acoustic Music Analysis methodology, is a relatively newly proposed approach to electro-acoustic music analysis. It was defined in a 2010 article by authors Tae Hong Park, David Hyman, Peter Leonard, and Wen Wu. The SQEMA “methodology is based on two main strategies: (a) exploitation of MIR[50] and salient feature extraction techniques and (b) employing a systematic analysis paradigm to segment a complex piece of music into smaller and more manageable parts.”[51] The group’s motivation in developing SQEMA stems from the lack of standardized electro-acoustic music analysis techniques and their research into MIR applications in electro-acoustic music analysis. Concerning the use of MIR and feature extraction techniques over an analysis methodology based on aural perception, the

---

[49] Ibid., 4.
[50] Music Information Retrieval, also sometimes referred to as Audio Information Retrieval or Audio Content Analysis when addressing feature extraction.
[51] Park, T. H., D. Hyman, P. Leonard, and W. Wu. “Systematic and Quantative Electro-Acoustic Music Analysis (Sqema).” (2010): 1.

authors assert that, “the innate subjectivity of perception and the current difficulty in accurately quantifying perceptual data places severe limitations on using only a perceptual approach for analysis and research.”[52] The authors go on to state that the “perceptual approaches hold value in the aesthetic evaluation and interpretation of a piece,”[53] and “they will hold even greater significance when backed by quantifiable data.”[54] The authors explain SQEMA with the use of EASY, a MIR/feature extraction application developed by the authors, which is not yet available to the public. However, many other MIR/feature extract environments exist: MARSYAS (well known C++ audio processing and MIR framework), MIRtoolbox (MATLAB® extraction library with segmentation and clustering tools), and jAudio (an easy to use portable interface but limited in functionality) among others.

In recent years there has been increasing interest and research in MIR in its various forms: audio feature extraction, classification, instrument recognition, mood and emotional classification, music aesthetics, and web applications among others. As the cost of computer processing power continually decreases, the capability to process ever larger amounts of data increases, allowing great strides to be made in the areas of audio feature extraction and other MIR fields not dealing in processing symbolic data (midi, music xml, xml4mir, meta data, and tags). Most MIR research has focused on traditional musics (those based on pitch, harmony, and rhythm), yet MIR is especially suited to electro-acoustic music’s focus on the evolution of timbre. There are many feature

[52] Ibid., 2.
[53] Ibid., 3.
[54] Ibid.

extraction algorithms that yield insight into the spectral content of a work, allowing measurable comparisons between materials as well as the ability to cluster materials and segment a work according to spectral content. Spectral content algorithms are perhaps the most useful in the analysis of electro-acoustic music, however, there are many other useful feature extraction options dealing with changes in amplitude, temporal regularity, pitch/harmonicity, and spatialization among others. Considering the diversity in electro-acoustic music, having a greater number of extraction algorithms with which to perform an analysis increases the probability of being able to plot the parameters significant to a given work. As might be expected, there is some amount of experimentation required in learning what feature extraction algorithms, or combination of algorithms, will yield the most insight in a given musical context. The following are descriptions of some algorithms useful in electro-acoustic music analysis:

> <ins>Brightness</ins> – a measurement of the energy present in high frequencies, useful in the scrutiny of timbre.
> <ins>Chromagram</ins> – measures the presence of pitch classes in an audio signal.
> <ins>Event Density</ins> – the rate of events, estimated by number of onsets per second, useful as a gauge for the general amount of activity in an audio signal.
> <ins>MFCC (Mel-Frequency Cepstral Coefficients)</ins> – a condensed description of the spectral content or timbre.
> <ins>Novelty</ins> – measurement of change in data, used to find points of change in sonograms, MFCCs, Spectral Centroid, or any other analysis performed.
> <ins>Sensory Dissonance</ins> – a measurement of the dissonance or "beating" phenomenon that occurs when pairs of sinusoids are close in frequency, estimated by

computing spectrum peaks, and taking the average dissonance between all possible pairs of peaks.[55]

Spectral Centroid – the frequency value at the center point/focus of the distributed spectral energy.

Spectral Flux – the measurement of change in spectral makeup, difference between the spectrums of successive analysis frames.

Spectral Spread – describes the concentration or standard deviation of the energy around the spectral centroid. [56]

Zero-crossing Rate/Noisiness – a method of estimating the noise content of a signal by counting the times the signal crosses the X-axis.

SQEMA's analysis procedure is a top-down model divide-and-conquer analysis paradigm organized into six steps: I – Multiple Listenings, II – High-Level Analysis, III – Mid-Level Analysis, IV – Event Level Analysis, V – Reexamination, and VI – Aesthetic Interpretation. The first step, or top-level, of the analysis is concerned with listening – the analyst will familiarize him/herself with the music through multiple listenings, at least starting without the use of visual aids. During this focused listening, judgments can be made as to how to guide the analysis steps to follow. Some of the tests of SQEMA have involved a listening technique called *companded listening*[57] – a listening technique developed by SQEMA's creators that "consists of compressing or expanding the duration of a piece with the option of doing so independent of its frequency content."[58] This

---

[55] Based on Sethares, 1998.
[56] Lerch, A. *An Introduction to Audio Content Analysis*. John Wiley & Sons, 2012. 47.
[57] Park, T. H., D. Hyman, P. Leonard, and P. Hermans. "Towards Comprehensive Framework for Electro-Acoustic Music Analysis." *ICMC* (2011): 2.
[58] Ibid.

listening technique allows "the listener to perceive the entire piece in a more 'manageable' period of time," [59] making some structural schemes easier to perceive. For the High-Level, Mid-Level, and Event Level Analyses (Levels II through IV), a "quantitative analysis is conducted via feature extraction techniques, clustering results from feature vector spaces, and observation of visualizations of extracted information."[60] The aim of the High-Level Analysis (II) phase is the segmentation of the work into its formal sections. The Mid-Level Analysis (III) phase focuses on the further segmentation of the formal section from phase II into *subsections*, *divisions*, and *events*. The authors define *divisions* as "segments that do not have sufficient saliency or contrasting features to be subsections,"[61] and *events* as "salient sonic occurrences that are hierarchically below the section, subsection, and division level."[62] The Event Level (IV) Analysis is concerned with event identification and applying descriptive labels informed by this most local focused analysis phase. The next phase of analysis, the Reexamination (V) phase, is concerned with cross-examination of all identified components from the formal section level to the event level. The connections between materials are identified and the design of connected materials is noted for trends. During the final phase of analysis, Aesthetic Interpretation (VI), the analyst uses the results of the previous analysis phases to interpret the work, rendering narrative and aesthetic elaboration of the work based on detailed quantitative information.

---

[59] Ibid., 2.
[60] Park, T. H., D. Hyman, P. Leonard, and W. Wu. "Systematic and Quantative Electro-Acoustic Music Analysis (Sqema)." (2010): 4.
[61] Ibid.
[62] Ibid.

The SQEMA methodology offers much for the electro-acoustic music analyst: a systematic approach for analysis, a means of confirming aural perceptions by working with empirical data, and a flexibility to incorporate other methodologies for analytical commentary. Where other methodologies have simply stated to start with segmenting the work, SQEMA provides a reasoned approach to accomplishing the segmentation process. Additionally, SQEMA negates the subjectivity of an analysis being based solely on listening.

It is unclear how much listening analysis is done in the segmentation process of an analysis. Is the role of listening simply to identify the qualities that would most clearly indicate divisions, or is a listening analysis performed and simply confirmed with MIR? The first option seems to lead toward a completely automated music analysis approach, while the latter is dependent on an adept analyst familiar with the difficulties involved in electro-acoustic music analysis.

3.5 – Methodology Conclusions

The methodologies detailed above are far from the only analytical approaches posited concerning electro-acoustic music, however, they represent a mix of prevailing and new approaches that are general in applicability and specific in focus. When combined into a single analytical approach, the presented methodologies can yield great insight into the nuances of a work. The strength and weakness of the listening based approaches is their focus, which makes them limited when used individually, but very effective if used in tandem. By combining approaches, a more complete analysis paradigm is created. SQEMA adds to such a paradigm with a logical process by which a work can be systematically scrutinized, and with a means of "checking" our ears via the quantifiable results of Music Information Retrieval.

The method of analysis is the first thing that must be evaluated in any analytical context. The detailed methodologies above will yield insight into all electro-acoustic works, but relying only on those analytical tools for all situations can lead to missed opportunities for understanding in the examination of some works. For example, it would be logical to make use of traditional analysis methodologies in the scrutiny of a work for acoustic instrument with electro-acoustic music accompaniment; the importance of pitch and rhythm must be addressed in such an analysis. Although not of importance in all electro-acoustic works, a discussion of source materials or possible processing techniques is sometimes of value. Spectromorphology is intended mostly for electro-acoustic music that is to a greater or lesser extent acousmatic, yet the identification of sound sources or *source-bonding* is acknowledged by Smalley as being of concern in the analysis of electro-acoustic music.

Chapter 4 - Analytical Case Study – Hiller's *Vocalise*

This chapter is an analytical case study, where the various issues related to electro-acoustic music analysis can be addressed through the process of an actual analysis. This chapter is titled an analytical case study and not an analysis because the following materials are not simply a presentation of the findings from an analysis, as would be the case in most documents of analysis. In this chapter, the steps taken to reach the results, the difficulties in the analysis, and observations are highlighted. The piece selected for the study, *Vocalise*, is the first movement from a larger work titled *Seven Electronic Studies for Two-Channel Tape Recorder* (1963) by the composer Lejaren Hiller (1924 - 1994). *Vocalise* was selected for a number of reasons: it has the appropriate scope and level of complexity for this case study; it is a work that has received relatively little exposure; and is a work that has yet to be analyzed or discussed by someone other than the composer. The recording used for this analysis is from the four disc set *In Celebration of the 50th Anniversary of the University of Illinois Experimental Music Studios (1958-2008).*

Hiller's contributions to electro-acoustic music are well documented (University of Illinois Experimental Music Studios studio reports, and *An Overview of the Music of Lejaren Hiller and an Examination of his Early Works Involving Technology* by James Bohn, among others). He founded the Experimental Music Studios at the University of Illinois at Urbana-Champaign (UIUC), one of the first electronic music studios in the United States, and contributed much in the field of computer-assisted algorithmic composition. Among the first works produced at UIUC's Experimental Music Studios, *Vocalise* seems the result of Hiller's intuition and not his experiments into computer-assisted algorithmic composition.

Vocalise is a compelling early example of electro-acoustic music. It was composed mostly with an instrument model (the sonic materials were presented/arranged as if to be performable by a person), without the amount of collage that most associate with electro-acoustic music. However, the work is not simple, but subtle, and has many interesting presentations/arrangements of materials. *Vocalise* makes use of both *concréte* and synthesized sounds, much like his other works of the time: *Nightmare Music* for Monarual Tape (1961); *Computer Contata* for Soprano, Chamber Ensemble, and Tape (1963); and *Machine Music* for Piano, Percussion, and Tape (1964).

Some insights into the work's construction can be gleaned from Hiller's program note from the disc booklet included with the four-disc set:

> In this study, the source of determinate pitches is twenty-four vowel formant peaks from the two resonant frequency regions associated with each vowel sound of ordinary spoken English. The structure of Study No. 1 [*Vocalise*] is organized according to the tripartite sectional form with the three sections labeled *Introduction and Statement, Development to Climax,* and *Vocal Fugue and Coda.*
> – Lejaren Hiller[63]

[63] In Celebration of the 50th Anniversary of the University of Illinois Experimental Music Studios (1958-2008). School of Music, University of Illinois at Urbana-Champaign. Print.

This program note gives insight into primary source materials and aspects of the formal organization, and while this information cannot be discounted in the analysis of the work, there is still much about the work that is left for an analysis to uncover.

4.1 – Analysis Method

The approach taken in this case study is modeled on the top-down SQEMA methodology, but makes use of other methodologies in the description of sonic characteristics and the functional role of materials. The analysis of Hiller's *Vocalise* is based on listening, and uses MIR as a tool for confirming the findings of the listening analysis. The MIRtoolbox[64] for the application MATLAB® – a very flexible environment for MIR feature extraction – is used for the MIR analysis. The study consists of the following sections: segmentation of the work, analytical commentary on the segmented materials, the identification of salient features, the analysis score, and a section for concluding statements.

The segmentation portion of the study consists of both listening and analysis via MIR. Appropriate points to sectionalize or chunk the work were located via listening, and followed up with MIR analysis to confirm measurable changes in the materials at those points. The listening aspect of the analysis informed which MIR processes to execute and examine. Some general visualization were generated (waveform, sonogram, etc.), so that observations could be made regarding what possible MIR process would be

[64] Olivier Lartillot, Petri Toiviainen and Tuomas Eerola developed MIRtoolbox as members of the Finnish Centre of Excellence in Interdisciplinary Music Research.

of value. The segmentation is divided into high level, mid level, and event level stages with the high level focused on identifying the primary sections of the work, the mid level focused on identifying sub-section and component layers of materials, and the event level focused on identifying significant events within the previously segmented materials.

The form, function, sonic characteristics, and similarities/differences of the segmented materials are discussed in the analytical commentary section. This section consists mostly of an interpretation and reexamination of the segmented materials, and makes use of further listening and examination of the MIR analyses as needed.

The section on salient features addresses the materials and ideas central to the work. This section also notes some of the audio processing and compositional techniques used in the work.

The analysis score section is simply a presentation of the analysis score in its entirety. The analysis score was created with the application EAnalysis – an application designed by Pierre Couprie for electro-acoustic music analysis.[65]

The conclusion addresses the difficulties involved with Hiller's *Vocalise*, a critique of the study, other observations, and a summary of future research based on the work done in this investigation.

---

[65] http://logiciels.pierrecouprie.fr/?page_id=402

4.2 – Segmentation

4.2.1 – Listening and First Observations

The segmentation process started with multiple listenings, as proscribed in SQEMA's top-most level of analysis. These listenings aided in the task of becoming familiar with the narrative of the piece, the materials used, and the general trajectories involved in how the piece unfolds. Three main sections are apparent and could be characterized by the simple presentation of a synthesized "vocal" material in the first section, a series of layered loops or repetitive events for the second section, and the heavy use of a delay effect on short vocal materials in the third section. Other observations at this time include:

- There are audible differences between the three main sections related to the frequency and manner in which events are presented: the relatively few events in the first section, give way to increased activity and cyclical rhythmic events in the second, and short, rhythmic bursts of audio with delays in the third.
- There is an economic use of source materials that includes synthesized and recorded sound: the synthesized "vocal" and noise materials; a few different recorded impacts, squeaks, and creaks; and various recorded vocal utterances.
- Concerning spatialization, there are either simple trajectories for the movement of sounds through both channels or the sound is positioned in one location in the stereo field.

Some general observations based on an examination of the waveform and sonogram of the work are listed here:

- There are some reoccurring contours present in Figure 10's Summed Waveform – three clear ramps. The first two having a short release after the peak of the ramp.
- The imbalance in channel energy in Figure 10's Stereo Waveform shows that the positioning of sounds occurs to the point of having some materials exist only in one channel. It is fruitful to examine each channel separately during the analysis process.

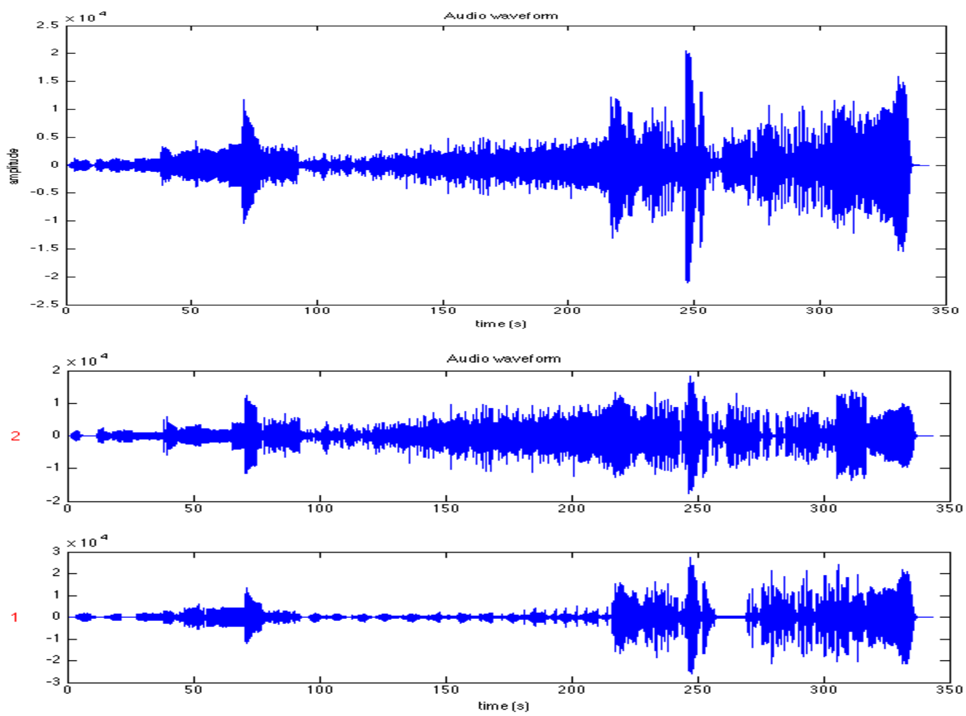


Figure 10 – [Top] Summed Waveform and [Bottom] Stereo Waveform display of *Vocalise*

- As illustrated in Figure 11, there is very little energy present in the work above 8000 Hz; specifying an upper limit frequency of 8000 Hz may optimize MIR analysis algorithms.

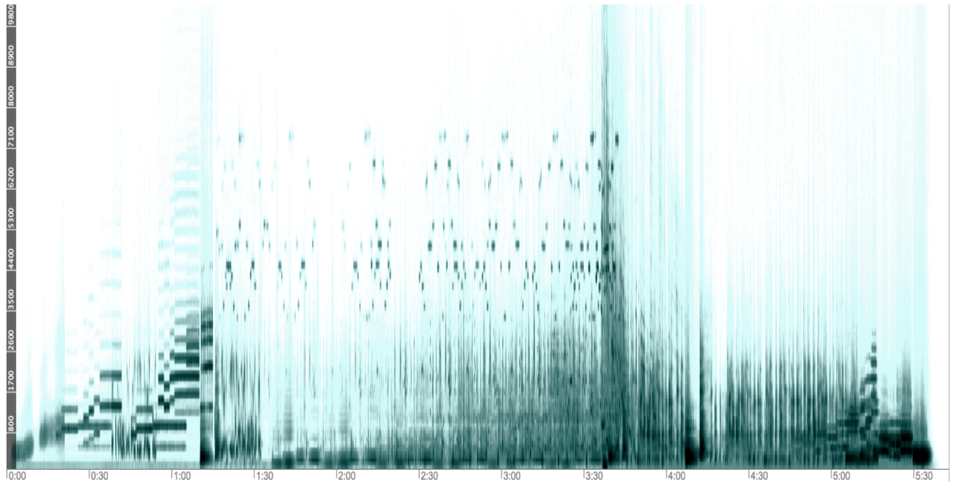


Figure 11 - Sonogram for *Vocalise* (log scale)

4.2.2 – High-level Analysis

Three main sections of the work were identified to be from 0 to 75 seconds, 75 to 229.1 seconds, and 229.1 to the end of the work (343 seconds). There are elisions of materials at both 75 seconds and 229.1 seconds, but these are the points at which the materials that characterize sections 2 and 3 respectively begin, and where the character of the piece shifts with the new materials.

To confirm these aural perceptions of the work are founded, first, the timbre summary of an MFCC analysis for the work was examined. Figure 12 is the generated visualization from the MFCC analysis, which displays very different activity from the beginning through just before 100 seconds than that of the rest of the piece. The

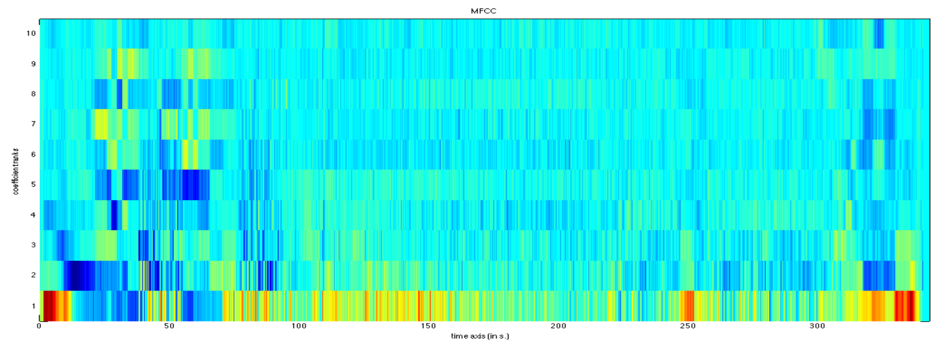


Figure 12 - MFCC Analysis of *Vocalise*

greater differences displayed in the materials between sections 1 and 2 are in line with the aural perceptions of the materials, and the area between 75 and 90 seconds in Figure 12 seems to confirm a continuation of some of sections 1's materials after 75 seconds. An examination of other timbre-based features should yield details of when the materials from section 1 cease and when section 2 materials begin. The differences between section 2 and 3 are much finer – by comparison, only slight differences in the trends of materials of section 2 and 3. For identifying clear trends around 229 seconds, an examination of timbre-based features will likely not be as useful as that of pitch or rhythm based features.

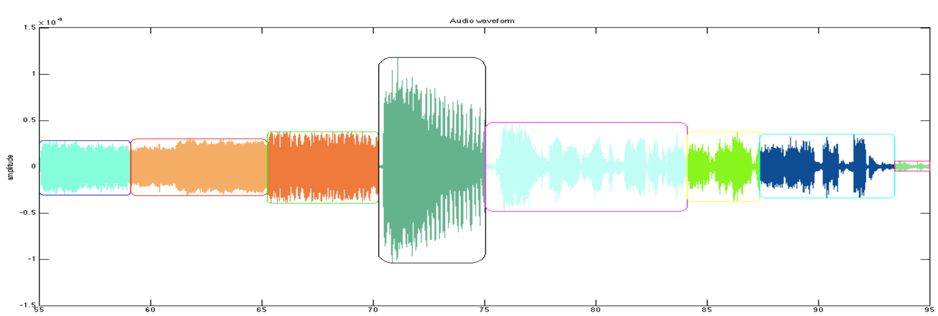


Figure 13 - *Vocalise* from 55 to 95 seconds Segmented by MFCC

The evaluation of timbre-based features around 75 seconds confirms very different trends in the make up of each of the surrounding section's timbres. For this analysis, the MFCC summary is used to first segment the work into chucks of like timbred material (Figure 13) before extracting the timbre-based features: spectral centroid, zero-crossing rate/noise estimation, and brightness for evaluation. As

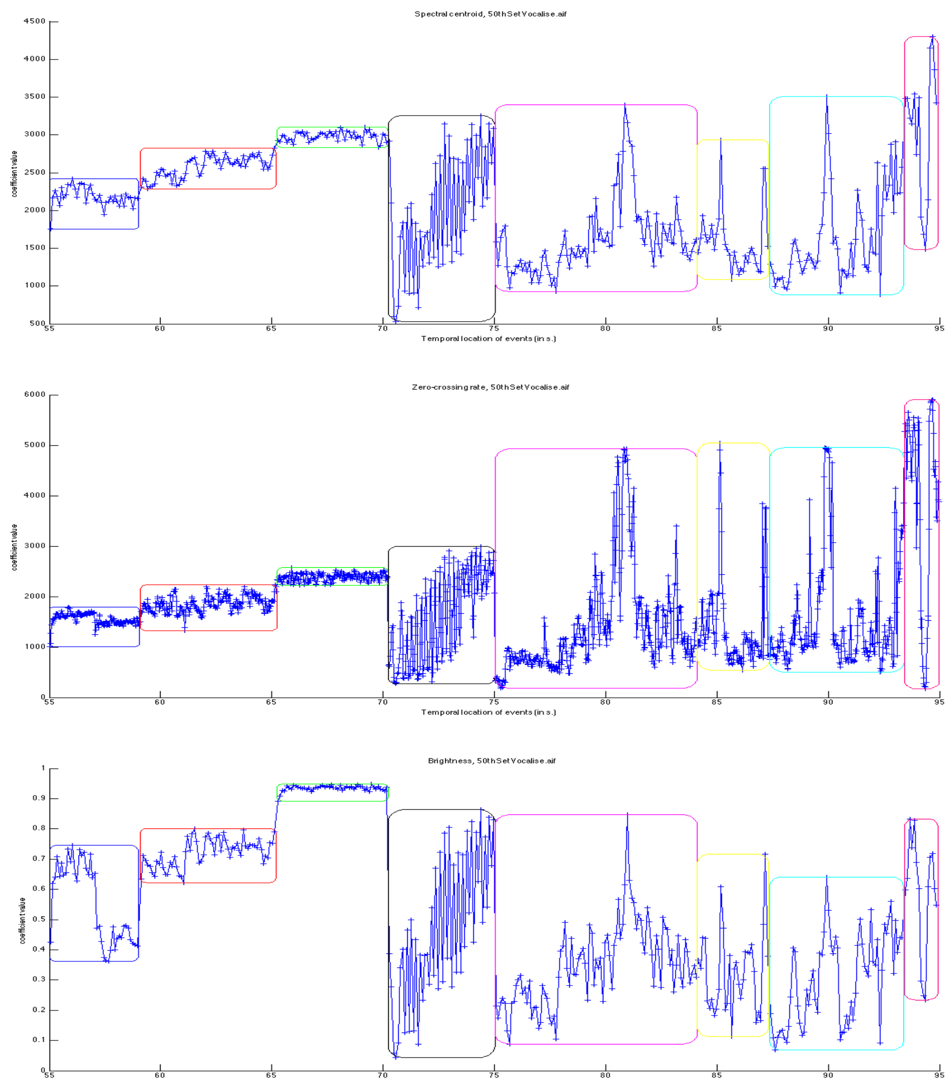


Figure 14 - [Top] Spectral Centroid, [Middle] Noise Content Estimation, and [Bottom] Brightness for *Vocalise* from 55 to 95 seconds with Segmentation

illustrated by the bounding boxes in Figure 13 and Figure 14, 75 seconds is a confirmed point of change of timbre in the work. The extracted features of Figure 14 illustrate the different trends of the materials before 75 seconds versus after 75 seconds. In general, the trend for the section 1 materials is that of terraced states. The exception being the last segment of section 1, from 70.2 to 75 seconds, which has a ramp contour in all three features graphed in Figure 14. The trend for section 2 is an ever changing and varied array of timbres, including frequent peaks and few steady state periods in the extracted features – quite different from the trends of section 1.

The examination of materials around 229 seconds was approached similarly, starting with segmenting from 200 to 260 seconds by the MFCC. The segmentation, displayed in Figure 15, confirms 229.1 seconds as a point of change of timbre in the work. Upon examination of the spectral centroid, noise content estimation, and brightness, the earlier assessment of the MFCC (Figure 12) was proven correct; there are much smaller differences in the trends of section 2 and 3 as compared to sections 1

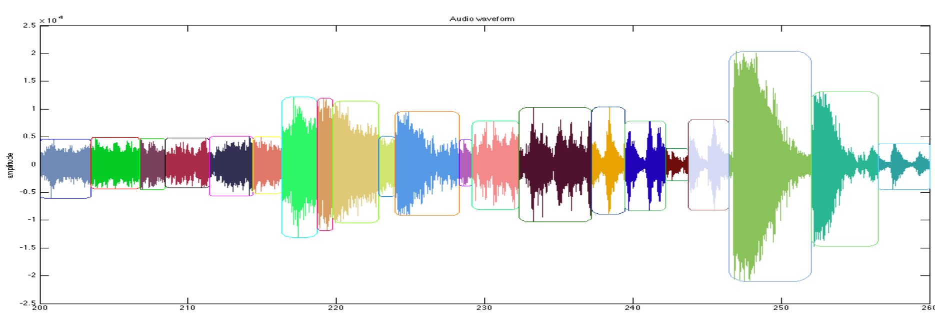


Figure 15 - *Vocalise* from 200 to 260 seconds Segmented by MFCC

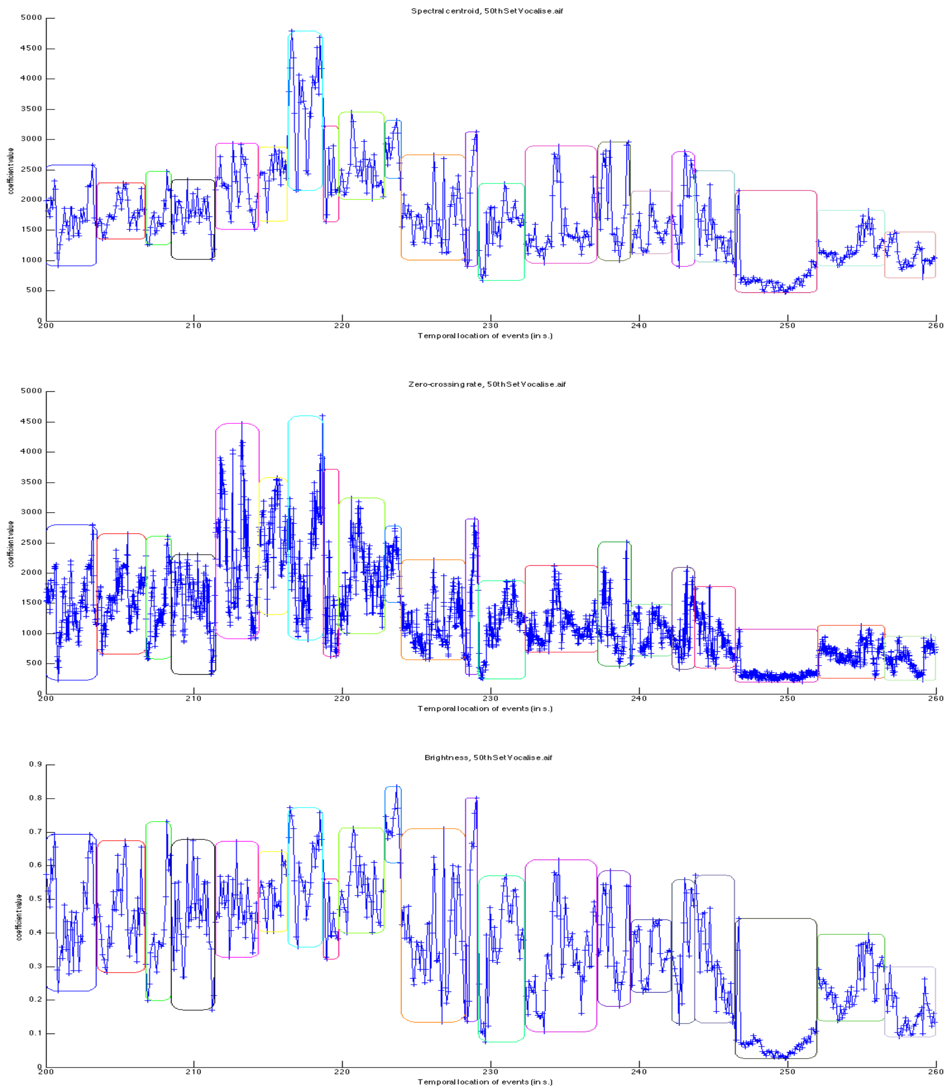


Figure 16 - [Top] Spectral Centroid, [Middle] Noise Content Estimation, [Bottom] Brightness for *Vocalise* from 200 to 260 seconds

and the general contour. The trend differences in Figure 16 are in the slight changes in bandwidth and descending contour before and after 229.1 seconds. The bandwidth changes are most evident in the widening of the spectral centroid and the narrowing of the noise content estimation. Seeking more contrast in the analysis data around 229.1, a pitch content analysis was performed. Figure 17 displays considerable differences in the

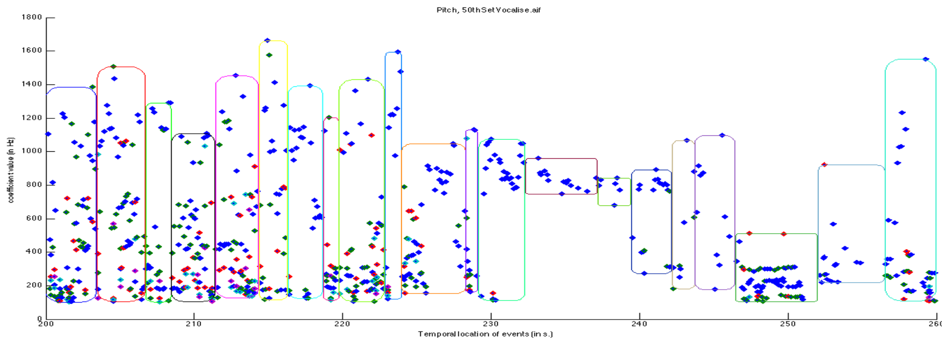


Figure 17 - Pitch Content for *Vocalise* from 200 to 260 seconds

trends of the work around 229.1 seconds. Section 2's pitch content is quite varied and has an average bandwidth of 3.5 octaves while section 3's pitch content is more contoured with an average bandwidth of 2 octaves.

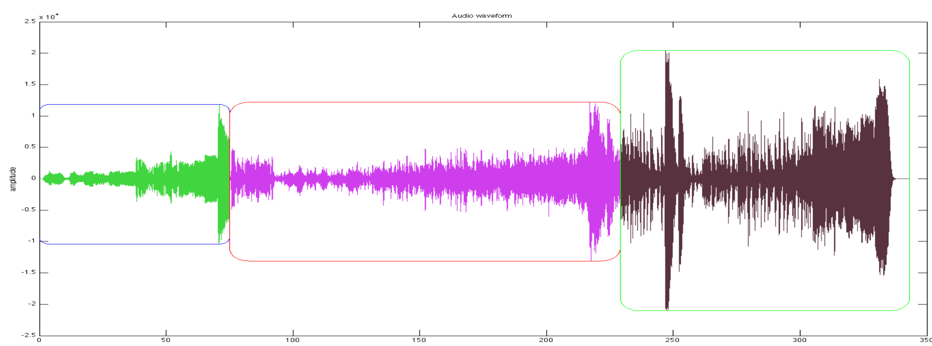


Figure 18 - *Vocalise* segmentation of main sections

4.2.3 – Mid-level Analysis

Segmentation at the mid-level consists of identifying the next level in formal division when it appears and/or the separation of the work into the top level of its component materials or layers.

4.2.3.1 – Section 1 [0 – 75 seconds]

Section 1 was aurally identified to contain four sub-sections with the divisions being at 17.7, 45, and 70.3 seconds. These sub-sections consist of a sub-section (0 – 17.7 seconds) of mid-to-background noise events, two sub-sections (17.7 – 45 and 45 – 70.3 seconds) of statement and restatement of the sections primary materials, and lastly a sub-section (70.3 – 75 seconds) of contrasting material to end section 1. The first sub-section consists of two 9 – 10 second events of mid-to-background noise, which are labeled Noise Event Type A. The two main sub-sections (17.7 – 45 and 45 – 70.3 seconds) consist primarily of a foreground presentation of synthesized pitched material with a vocal quality, which is labeled Vocal Synth Material. The Vocal Synth material is layered from 38 – 54 seconds overlapping the section division of 45 seconds with short bursts of another noise material, which is labeled Noise Event Type B. The final sub-section consists of both noise and pitched materials presented differently than those so far in the work. There is a single noise event that occurs over the duration of the sub-section, which is labeled Recorded Noise Event A, and some pitched material that presents a sonority for the length of the sub-section in a repeated ramping fashion.

To confirm the aural analysis, the section was first segmented according to the MFCC, and finding that the results were not as detailed as needed, was subsequently segmented by the spectrum. Figure 19 shows the segmented results and the novelty graph of the peaks of spectral change in the section. This type of segmentation shows

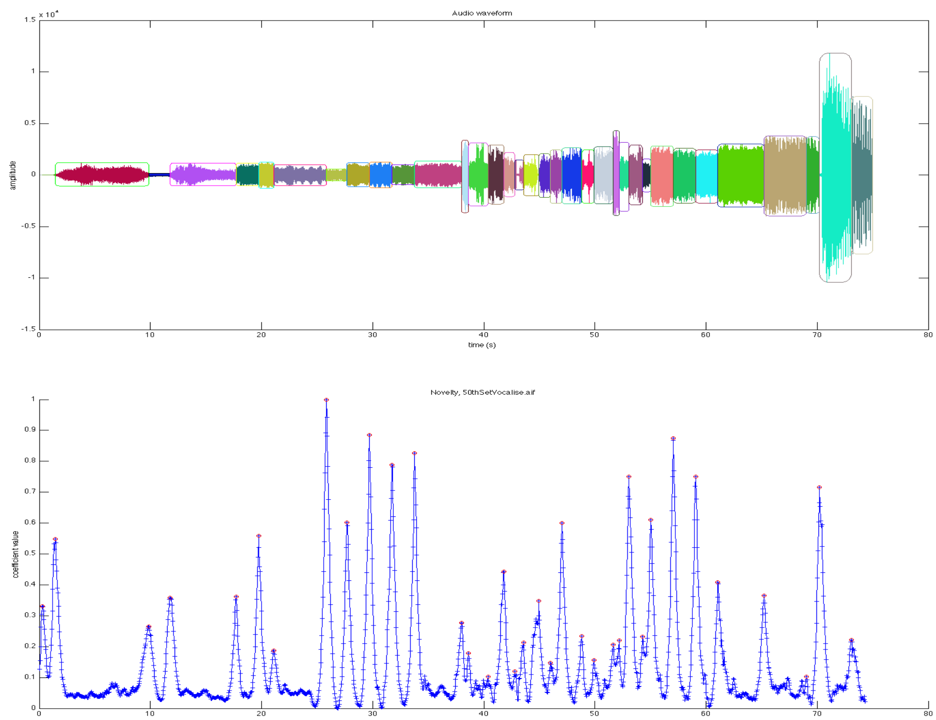


Figure 19 - [Top] Section 1 Segmentation according to peaks of [Bottom] Spectral Novelty Graph

most of the events as points of division; even re-articulations of the same type of event will show small peaks on the novelty graph as in evidence around 45 seconds with the Noise Event Type B material. Timbre based feature extraction yielded little useful information in the analysis at this level. To differentiate between the two primary material types of the section, noise-based and pitched sounds, a chromagram was utilized to visualize when the full chromatic, or noise, and when only a few pitch classes are present. The arrangement and concentration of pitch classes in the chromagram, Figure 20, confirms the arrangement of materials as explained above: noise sub-section at the start of the work consisting of two noise events at 1.4 and 11.8 seconds;

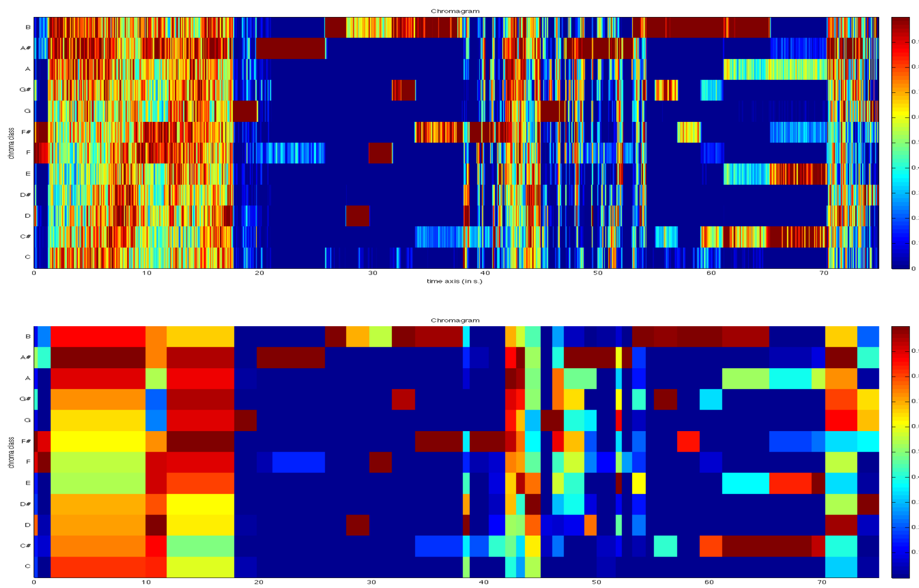


Figure 20 - [Top] Section 1 Chromagram and [Bottom] Chromagram by the Novelty Segmentation from Figure 19

two similar primary sub-sections starting at 17.7 and 45 seconds with the starting pitch classes G, A#, and B; noise-based layer of material between the primary pitched statements; and a closing sub-section with both noise and pitched materials.

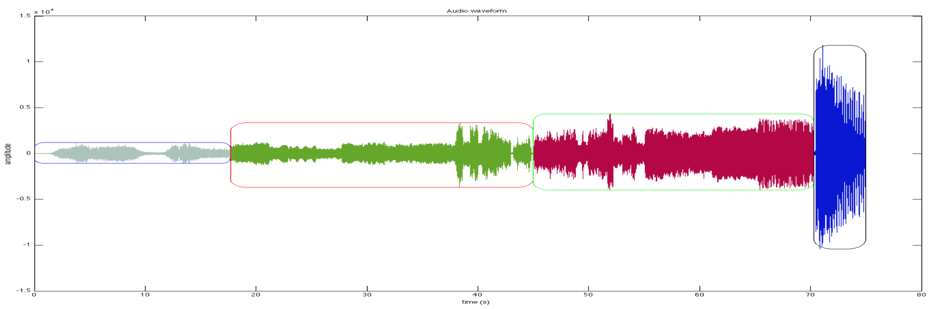


Figure 21 - Section 1 Formal Divisions

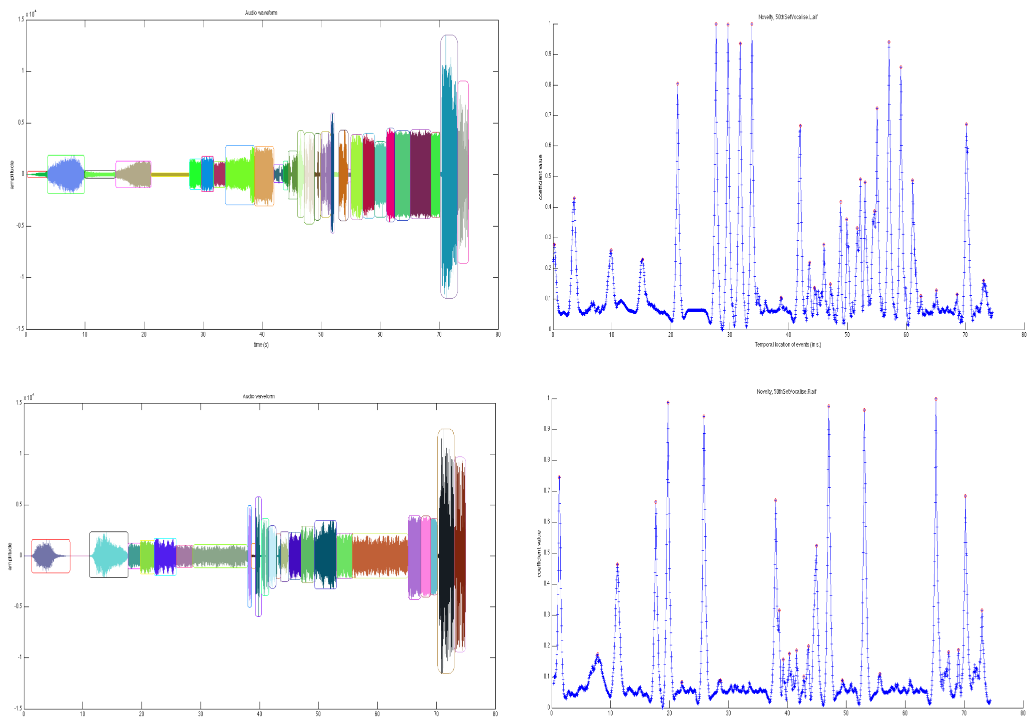


Figure 22 - [Top Left] Section 1 Left Channel Segmentation according to peaks of [Top Right] Spectral Novelty Graph and [Bottom Left] Section 1 Right Channel Segmentation according to peaks of [Bottom Right] Spectral Novelty Graph

Section 1 can also be easily separated into its component layers; for the duration of Section 1, there is a considerable amount of separation between the left and right channels as demonstrated by the differences in segmentation and spectral novelty graphs in Figure 22. In general, the left channel is delayed from the Right Channel until the closing sub-section of Section 1 at 70.3 seconds:

- Sub-section 1 (0 – 17.7 seconds); both Noise Event Type A occurrences start in the right channel before panning to the left.
- The second Noise Type A Event in sub-section 1 ends at 17.7 in the right channel but at 21.2 in the left channel.
- In the statement and restatement of the section's primary materials sub-sections, the right channel starts the section with the G, A#, and B pitch classes while the left channel has a delayed second stream of pitch classes that are different for each sub-section.
- Lastly, the emergence of the Noise Event Type B material again is presented in the right channel (38 seconds) before it occurs in the left channel ~3 seconds later and the material ends similarly delayed with the right channel at 45 seconds and the left channel at 55 seconds.

This channel separation can be visualized in the Segmented Chromagrams of Figure 23. The result of this channel separation is overlapping layers of contrasting materials: 0 – 21.2 seconds, Noise Type A Events layer; 17.7 – 42 seconds, pitched materials layer; 38 – 54.4 seconds, Noise Type B Events layer; and 45 – 70.3 seconds, pitched materials layer restatement. The areas in Figure 23 in sub-section 1 that show an emphasis of the pitch class F# are due to the normalization of the visualization. The F# here is the dominant pitch of the system hum of the audio.

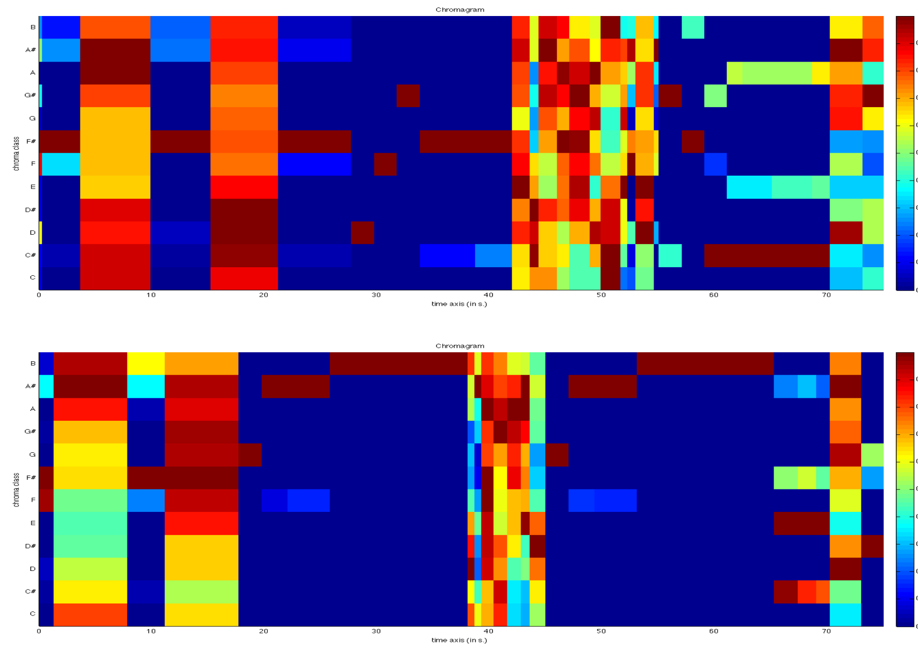


Figure 23 - Section 1 [Top] Left and [Bottom] Right Channel Chromagram by the Novelty Segmentation from Figure 222

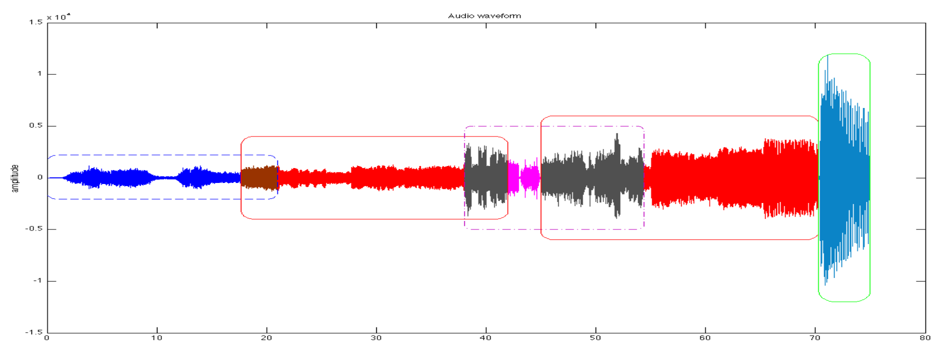


Figure 24 – Section 1 layers Segmentation

4.2.3.2 – Section 2 [75 – 229.1 seconds]

Section 2 has three aurally identifiable sub-sections: 75 – 93.47 seconds, 93.47 – 216.2 seconds, and 216.2 – 229.1 seconds. Sub-section 1 can be characterized by the elision of Section 1 and 2 materials. In sub-section 1, there are Recorded Noise Event A and Noise Event Type B materials, as well as two new materials: a new presentation of the Vocal Synth material (transposed and possibly otherwise varied) and a *concréte* (or pre-recorded sound based) layer of material with mostly percussive sounds. Sub-section 2 is a building section-starting with minimal activity and progressively becoming more complex with the continuous addition of mostly looped or repetition based layers of materials. Sub-section 3 is the apex of complexity for the work and introduces yet more new materials, however, in an overall trajectory toward simpler activity by the end of the sub-section.

With the increased complexity of Section 2, seeking confirmation via MIR for the above sub-section decisions becomes less straightforward than in Section 1. The

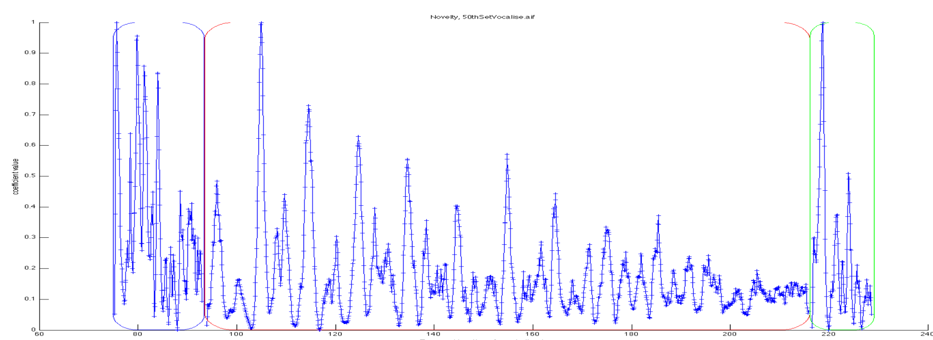


Figure 25 - Section 2 Spectral Novelty Graph

increased complexity of Section 2 makes it more difficult to identify new trends, as there are often separate trends for each layer of materials; the complexity translates to the MIR feature extraction. For example (Figure 25), the trending of novelty in Section 2 does show slight separation of trends at 93.47 and 216.2 seconds, but it is not clear enough for it to be the only analysis performed on this section. To combat this complexity, as with

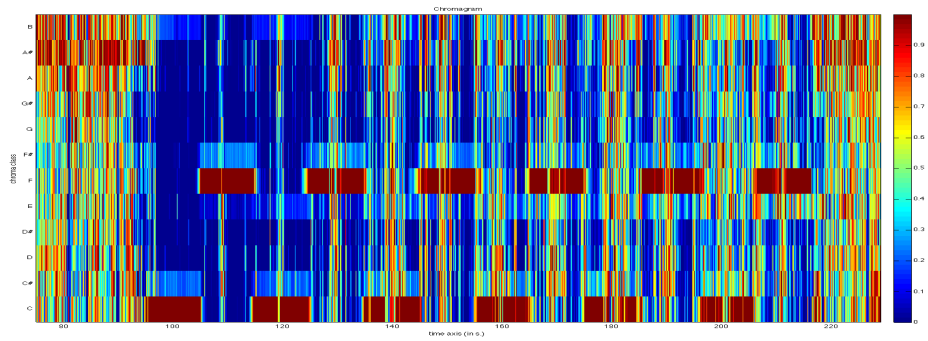


Figure 26 - Section 2 Chromagram (wrapped to 1 octave)

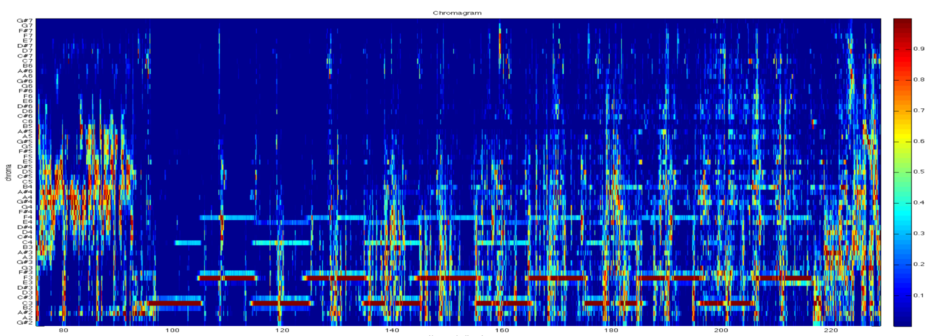


Figure 27 - Section 2 [Top] Chromagram (unwrapped)

Section 1, the focus is on how each sub-section is different, to see if any trends can be identified based on aspects unique to the specific materials. Through deductive reasoning and some degree of trial and error, confirming analysis results can be achieved: Figure 26 displays prominent alternating C and F pitch classes that only occur in sub-section 2 (93.47 – 216.2 seconds), as well as the presents of the full chromatic in sub-sections 1

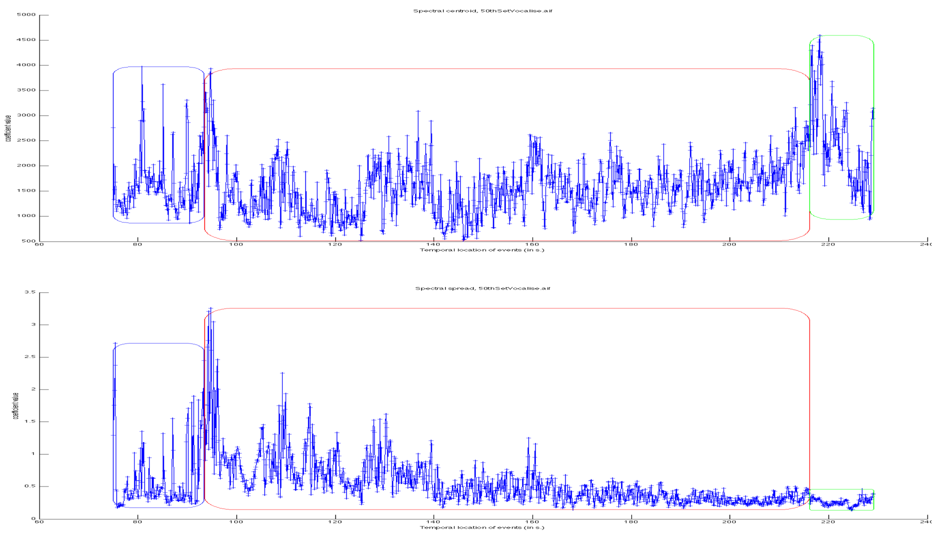


Figure 28 - Section 2 [Left] Spectral Centroid and [Right] Spectral Spread Graphs

and 3, and the increasing presences of the full chromatic in sub-section 2; and Figure 27 confirms the same with the addition of how the concentrations of materials in sub-section 1 and 3 are arranged. For more details on the spectral makeup of the three sub-sections, a

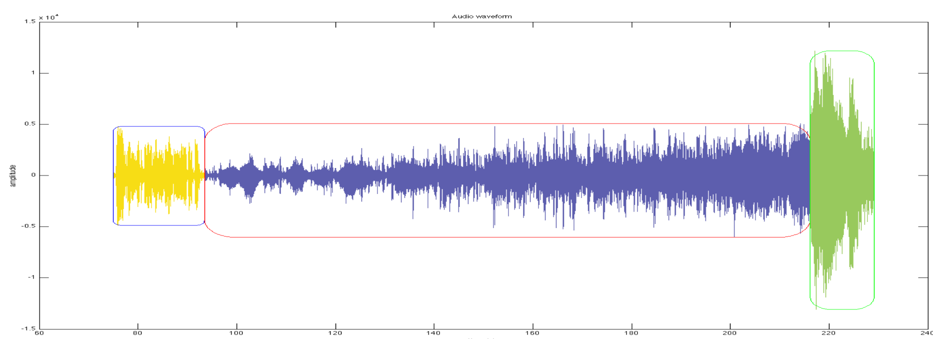


Figure 29 - Section 2 segmentation of formal divisions

spectral centroid and spectral spread analysis (Figure 28) were performed. This further highlighted the parameter contours and differences of Section 2's sub-sections and how they unfold.

The complexity of Section 2 is reached through the use of many layers of materials. Most of the layers were easily identified after examining the channels separately, however, some required the use of a spectrogram and many listenings before they could be identified with confidence. The material layers for Section 2 are as follows:

- Starting in sub-section 1:
    - 75.5 – 92.87 seconds, a Noise Event Type B layer in both channels
    - 76.1 – 113.5 seconds, a higher (above 3000 hz.), very low amplitude, and perhaps modified version of the Vocal Synth used in Section 1 with new pitch content in the right channel
    - 78.5 – 222 seconds, a layer of *concréte* materials consisting mostly of a ~16.5 second loop, but with some other materials added in toward the end (*Concréte* Materials/Loop) in the right channel; the recorded materials include a prominent door squeak/creaking recording and other sounds more percussive in nature

- Starting in sub-section 2:
    - 93.47 – 151 seconds, a repetitive new noise event type layer (Noise Event Type C) in the left channel
    - 94.8 – 216.25 seconds, a "bass line" type material that alternates between pitch classes C and F as well as left and right channels
    - 100.5 – 222.6 seconds, a repetitive ramping event layer in the left channel
    - 121.9 – 139.8 seconds, more transposed Vocal Synth material in the right channel

- o 149.9 – 223.1 seconds, more transposed Vocal Synth material in the right channel

- Starting in sub-section 3:
    - o 216.2 – 218.75 seconds, repetitive attacking bass gesture with slight upward then downward glissandi in the left channel
    - o 216.2 – 248.75 seconds, a *Concréte* Based Stuttering or re-attacking material with a downward glissandi at multiple strata in the left channel (elides with Section 3)
    - o 217.6 – 229.1 seconds, similar to the last layer but in the right channel and with a lower amplitude

As with seeking confirmation of the formal divisions of Section 2, the increased complexity here makes differentiating between the separate layers of materials more difficult. In some cases, this required the use of filtering in addition to the examination of separate audio channels to isolate specific frequency ranges so as to focus on materials otherwise overshadowed by more prominent layers.

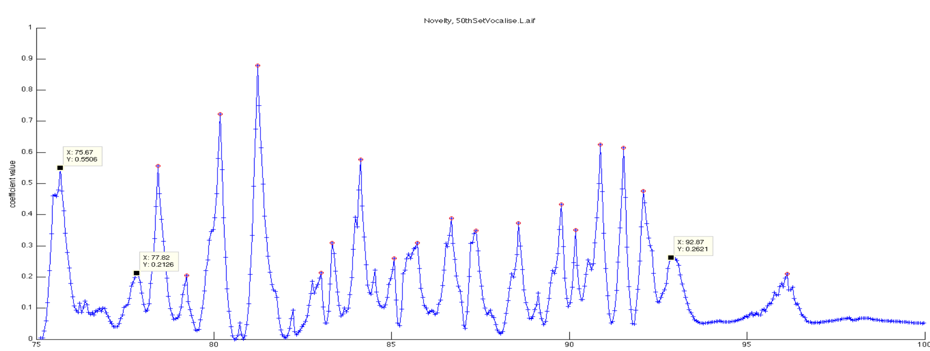


Figure 30 - Section 2 sub-section 1 left channel Spectral Novelty Graph

Starting in sub-section 1, Figure 30 displays the left channel spikes in spectral change that indicate new events, or moments of silence as is the case of some of the higher peaks here. The labeled peaks show the starting of new trends in spectral activity in the left channel: 75.67 seconds, the start of a modified Recorded Noise Event A; 77.82 seconds, the start of the Noise Event Type B layer; and 92.87 seconds, the start of sub-section 2's repetitive Noise Event Type C layer in the left channel. Figure 31 further

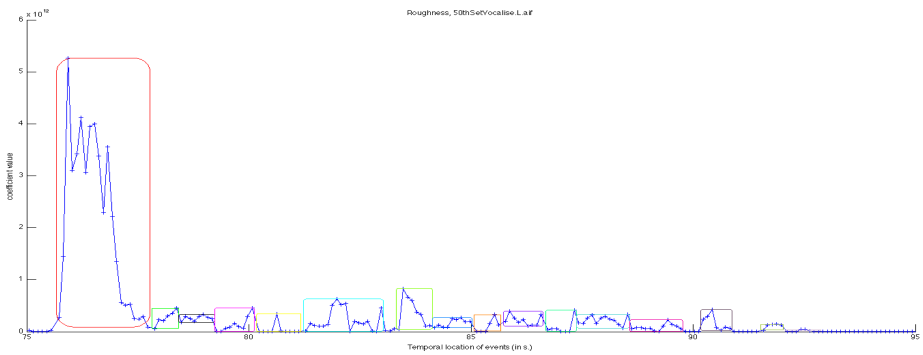


Figure 31 - Section 2 sub-section 1 left channel Sensory Dissonance Graph

highlights the differences in the layers here, indicating three very different dissonance profiles for the layers noted in Figure 30. In the right channel, the difficulty in isolating the materials of each layer was considerable. Two of the layers occur in the same frequency range with similar sonic characteristics and the last layer occurs at a very low amplitude with similar issues of materials from other layers obscuring the analysis. The Noise Event Type B layer and the *concréte* layer overlap from 78.5 to 92.87 seconds, and considering the *concréte* layer at this time has a fairly low amplitude level, it is

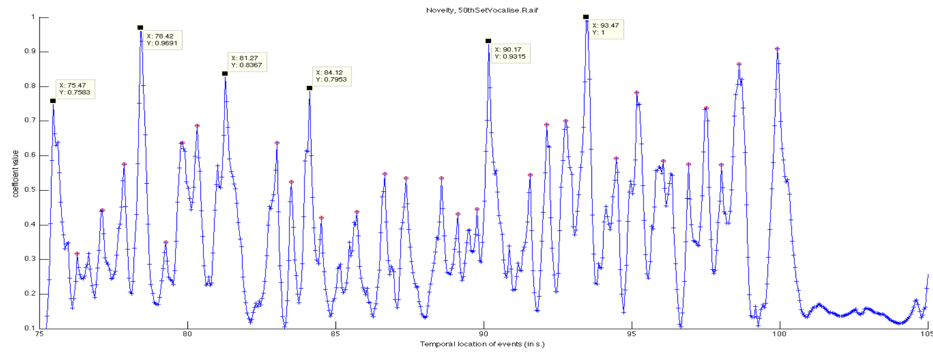


Figure 32 – Section 2 sub-section 1 right channel Spectral Novelty Graph

understandable that the layers would be difficult to differentiate via MIR. This is true of the percussive recordings that make up much of the *concréte* loop at this point, however, the door squeak/creaking sounds are quite different in sonic character than the percussive recordings and Noise Event Type B materials, and therefore are able to be isolated with some confidence. Figure 32 displays the spikes in spectral change for the right channel of sub-section 1 including the start of the Noise Event Type B layer ~75.5 seconds as well as the first three occurrences of the door squeak/creaking sound of the *concréte* loop layer (78.42, 81.27, and 84.12 seconds.) The other labeled peaks are moments of abrupt silence in the Noise Event Type B layer of materials. A comparison of before and after the end of the Noise Event Type B layer (92.87 seconds) in Figure 33's analysis results

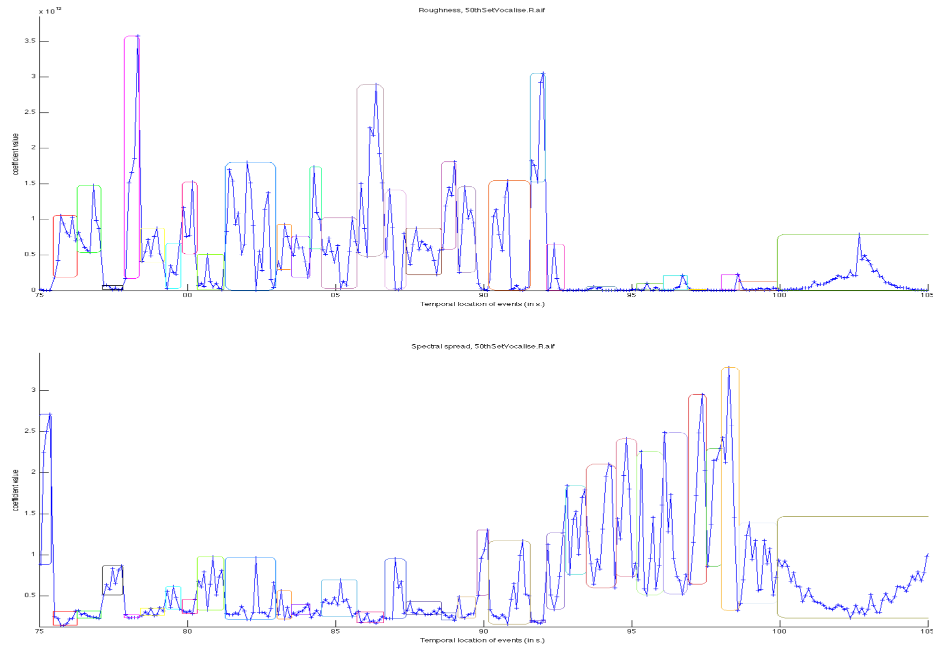


Figure 33 - Section 2 sub-section 1 right channel [Top] Sensory Dissonance and [Bottom] Spectral Spread Graphs

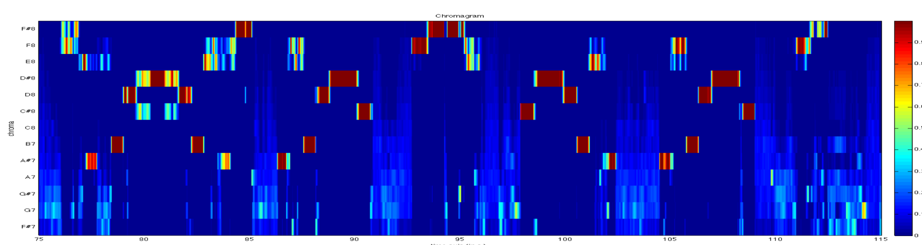


Figure 34 - Section 2 sub-section 1 Chromagram after filtering and normalization

clearly illustrates the prominence of the Noise Event Type B layer in sub-section 1. Due to the minimal amplitude level, the frequencies below 3000 Hz. were filtered out to

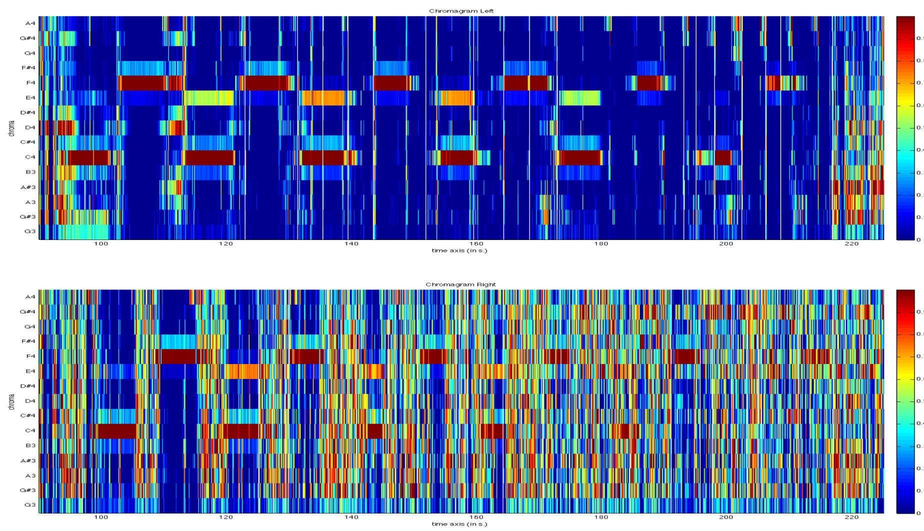


Figure 35 - Section 2 sub-section 2 Chromagram [Top] Left Channel and [Bottom] Right Channel

isolate the first instance of the new Vocal Synth-like material (76.1 – 113.5 seconds). To have audio useful for analysis, the filtered result was normalized. The chromagram below, Figure 34, illustrates the bounds of this layer of material, as well as the symmetry of its construction.

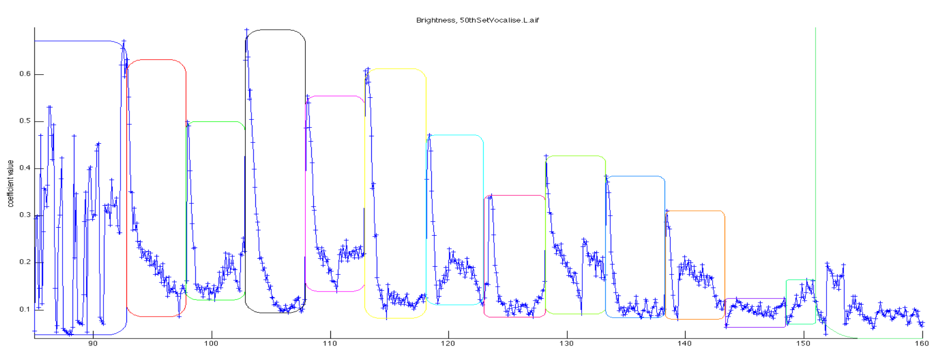


Figure 36 - Section 2 sub-section 2 right channel segmented Brightness Graph

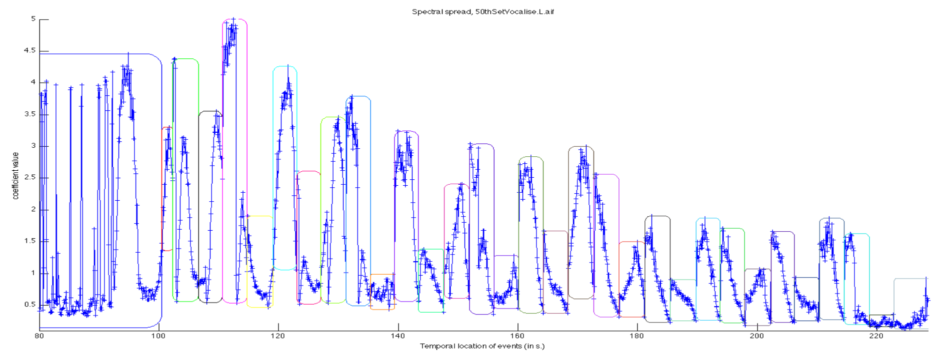


Figure 37 - Section 2 sub-section 2 filtered left channel Spectral Spread Graph

In sub-section 2, the layers are, for the majority, similarly separated into left and right channels. The exception to this is the "bass line" like layer that occurs from 94.8 – 216.25 seconds; its bounds can be identified clearly in Figure 26 and Figure 27. Figure 35's channel separation provides less clarity for this layer's bounds than insight into the overlapping and delayed arrangement of the layer's materials, as each pitch begins in the left channel before appearing in the right channel. For the duration of sub-section 2, excluding the "bass line" layer, the Noise Event Type C and repetitive ramp layers occupy the left channel. The Noise Event Type C layer can be most easily confirmed by the brightness contributed by its repetitive attacks. Segmented according to the attacks of this layer, Figure 36 clearly illustrates the characteristic shape for the attacks of the Noise Event Type C layer that is mostly absent before 93.47 seconds and after 151 seconds. The ramp material layer is most easily identified by the abrupt termination at the end of each ramp. The segmentation in Figure 37 is from each ramp termination to the next ramp termination; the contour of spectral spread toward the termination of each ramp characteristically plummets making the bounds of the layer able to be

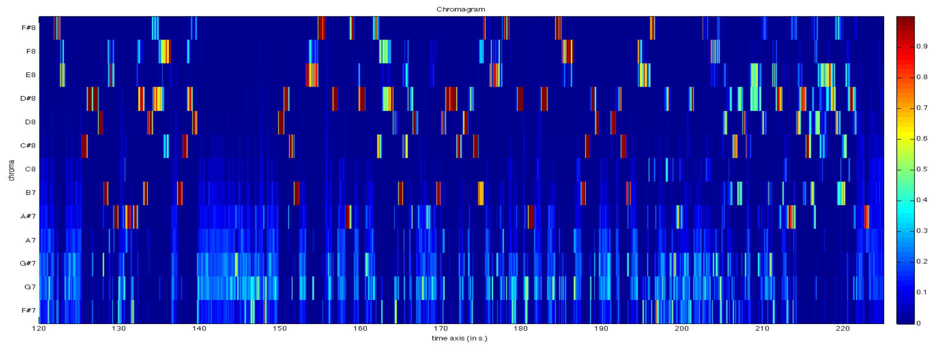


Figure 38 - Section 2 sub-section 2 filtered and normalized Chromagram

identified with confidence. In the right channel, there is the high Vocal Synth-like material continuing from sub-section 1 until 223.1 seconds, albeit with some breaks. The first break is illustrated in Figure 34 (113.5 seconds); the other break in this material is

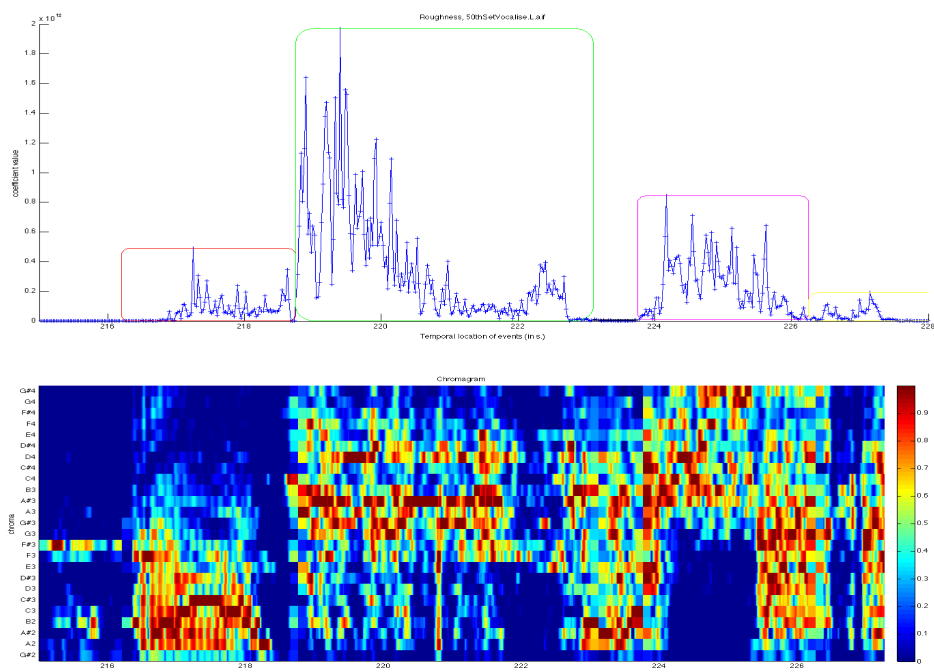


Figure 39 - Section 2 sub-section 3 filtered left channel [Top] Sensory Dissonance Graph and [Bottom] Chromagram

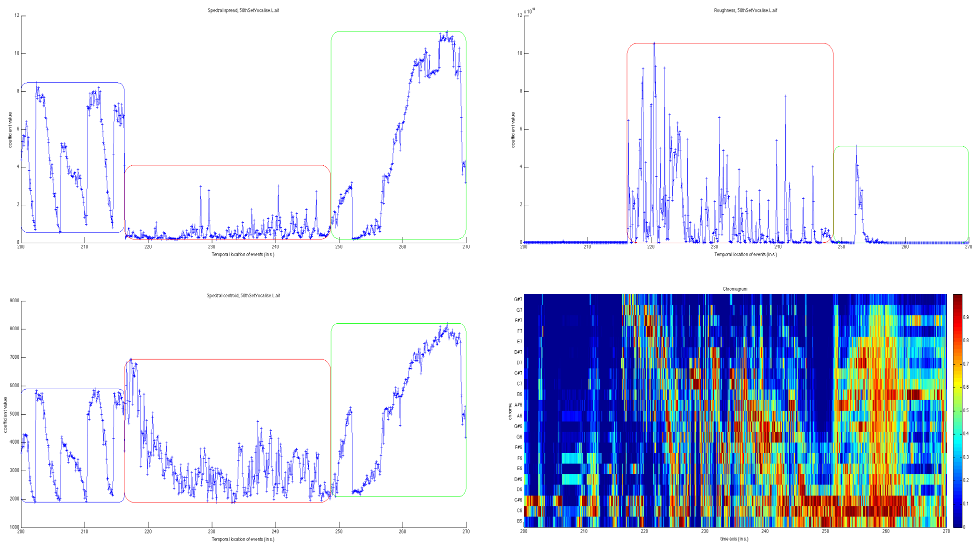


Figure 40 – 200 to 270 seconds filtered left channel [Top Left] Spectral Spread, [Top Right] Sensory Dissonance, [Bottom Left] Spectral Centroid, and [Bottom Right] Chromagram analysis results

confirmed in Figure 38 to be from 139.8 – 149.9 seconds with the layer finally terminating at 223.1 seconds.

In sub-section 3, there is still a great deal of separation between the activities of the left and right channels. Occurring in the left channel, the repetitive attacking bass gesture (216.6 – 218.75 seconds) is well represented by Figure 39, which illustrates the sensory dissonance profile and glissandi aspect of the gesture, as well as the following Recorded Noise Event A and Recorded Noise Event B. Due to the relative strength of the *Concréte* Based Stuttering material, Figure 39 required filtering out frequencies above 800 Hz. to achieve a meaningful analysis result. The *Concréte* Based Stuttering layer occurs in the left channel from 216.2 – 248.75 seconds and is the primary sonic material of sub-section 3. The analysis of frequencies above 800 Hz., Figure 40, illustrates its

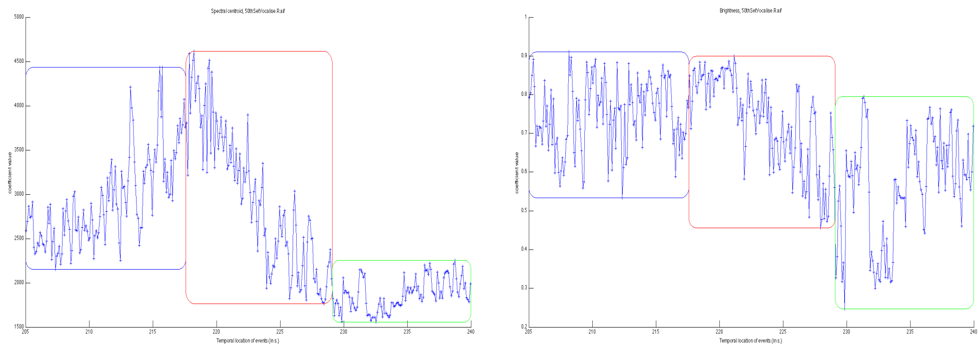


Figure 41 - 205 to 240 seconds filtered right channel [Left] Spectral Centroid and [Right] Spectral Brightness Graphs

contrasting timbral makeup as compared to its surrounding materials as well as its characteristic downward trajectory in the spectral centroid and chromagram. Similar contours can be seen in the right channel, Figure 41, for the occurrence of the *Concréte* Based Stuttering material from 217.6 to 229.1 seconds. Lastly, in the right channel, the *concréte* layer continues from sub-section 1 (78.5 seconds) to 222 seconds.

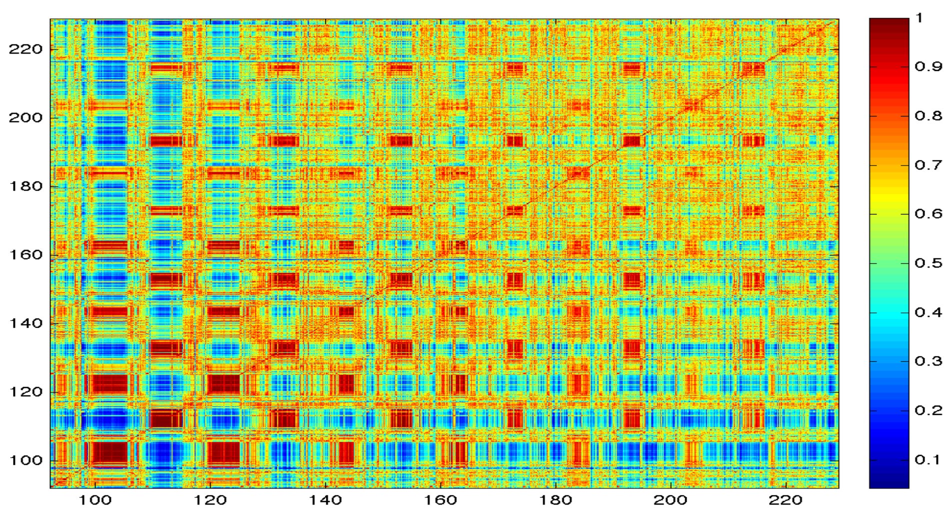


Figure 42 - Section 2 sub-section 2/3 filtered right channel Spectral Similarity Matrix

Unfortunately, confirming its termination is very difficult due to the presence of the *Concréte* Based Stuttering material, which shares a significant amount of its timbral makeup with the c*oncréte* loop layer. Figure 42's similarity matrix confirms the increasing similarity in timbral makeup throughout Section 2. The last events of the *concréte* loop layer were identifiable, but the analysis results hardly shows their passing as significant. Figure 43 illustrates the spectral novelty of the last two events of the layer as well as the pitch makeup: the first, a glancing F3 to G3 event and the final event simply a G3 echo of the previous event.

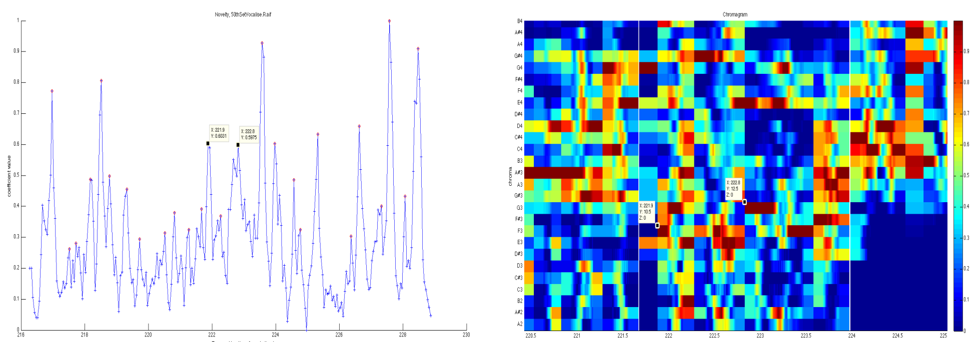


Figure 43 - Section 2 sub-section 3 filtered right channel [Left] Spectral Novelty Graph and [Right] Chromagram

4.2.3.3 – Section 3 [229.1 – 343 seconds]

Section 3 can be aurally identified to consist of four sub-sections: 229.1 – 246.6 seconds, 246.6 – 255.8 seconds, 255.8 – 310.4 seconds, and 310.4 – the end of the piece (343 seconds.) The last section of the work can be characterized by the emergence of what appears to be short vocal utterances processed with delay as the primary sonic material. The first sub-section (229.1 – 246.6 seconds) is the first appearance of this characteristic Vocal Delay material. This sub-section consists of the Vocal Delay material (dry) and the *Concréte* Based Stuttering material from the last sub-section of Section 2 in one channel, and reverb processed Vocal Delay material in the other channel. Sub-section 2 (246.6 – 255.8 seconds) is a short section made up of two processed recorded/*concréte* based events. In sub-section 3 (255.8 – 310.4 seconds), the characteristic Vocal Delay material returns in an imitative fashion, with the initial statement in the right channel and the answer in the left channel ~15 seconds later. The final sub-section of work (310.4 – 343 seconds) is a return of the Vocal Synth material similar to that at the end of Section 1 sub-section 3, with some additional low frequency noise elements.

To confirm this analysis via MIR, the peaks of a spectral novelty analysis were first examined. Finding that the times 246.6, 255.8, and 310.4 seconds are in fact points of spectral change in Section 3 (Figure 44), several of the timbre based MIR analyses

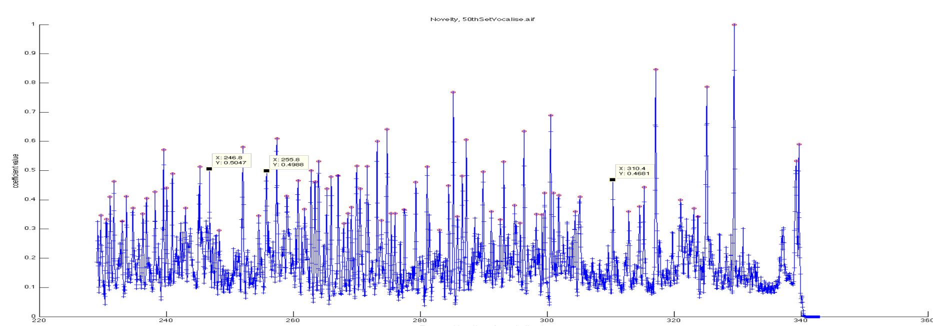


Figure 44 - Section 3 Spectral Novelty Graph

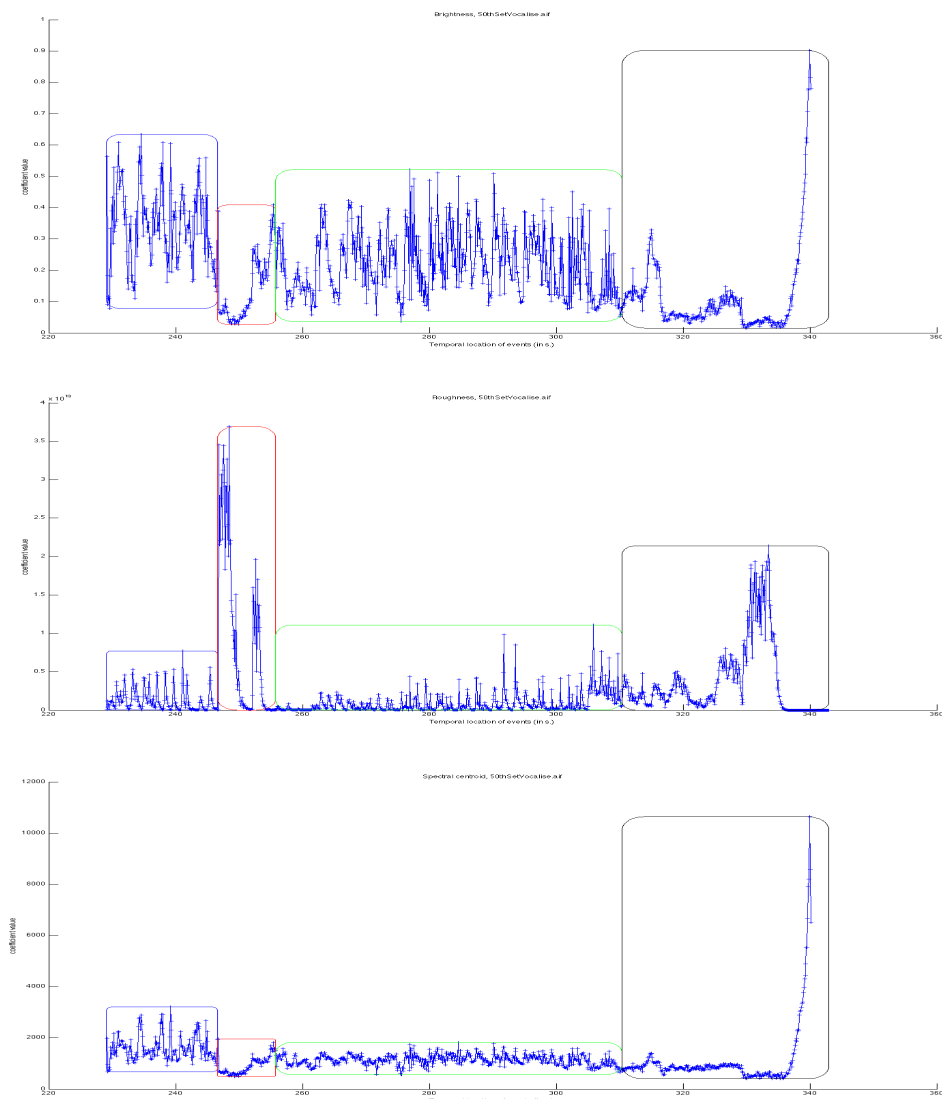


Figure 45 - Section 3 [Top] Spectral Brightness, [Middle] Sensory Dissonance, and [Bottom] Spectral Centroid Graphs

were performed. The above times did not prove to be the most profound spectral changes of the section, yet Figure 45 confirms that they are the points at which new trajectories commence in timbral makeup. Figure 45 also confirms the similarity of sub-sections 1 and 3, as well as the contrasting makeup of sub-sections 2 and 4 as compared to 1 and 3.

As with Sections 1 and 2, Section 3 maintains a fair degree of channel separation and layering of materials. This is evident in the first sub-section, as explained above briefly, where the left channel continues the *Concréte* Based Stutter layer of material from the previous section until 248.75 seconds and the first appearance of the Vocal Delay material is present from 229.1 to 246.6 seconds, while in the right channel a separate layer of the processed Vocal Delay material is also presented from 229.1 to 246.6 seconds. Layering occurs again in sub-section 3: the right channel presents the Vocal Delay material from 255.8 to 310.4 seconds while the left channel presents a separate stream of the Vocal Delay material only from 269.9 to 310.4 seconds. Lastly, the Vocal Synth Plus (“Plus” due to the occasional support by low frequency noise elements) material that closes the work emerges well before the start of sub-section 4 (285.2 seconds) – layered with the Vocal Delay material.

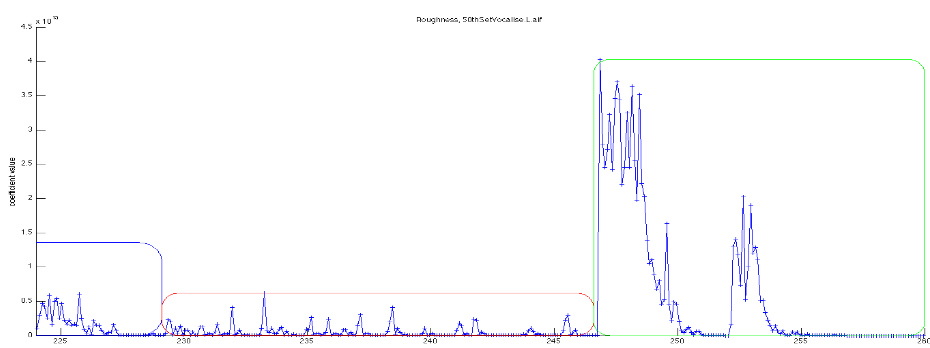


Figure 46 - Section 3 sub-section 1 filtered left channel Sensory Dissonance Graph

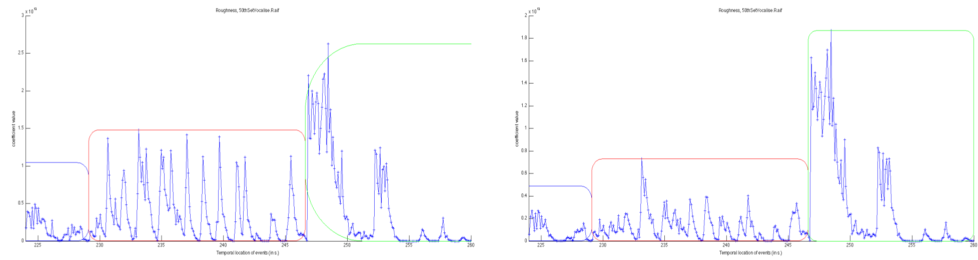


Figure 47 - Section 3 sub-section 1 right channel [Left] unfiltered and [Right] filtered Sensory Dissonance Graphs

Confirmation of the presence of the separate layers of material in Section 3, as in Section 2, is much more difficult than confirming formal divisions. Seeking confirmation for the Vocal Delay layer in the left channel from 229.1 to 246.6 seconds, the majority of the frequency range of the *Concréte* Based Stutter material from Section 2 was first filtered out – a 1000 Hz lowpass filter. Figure 46 illustrates the differences in the sensory dissonance trends of this layer from its neighboring materials. As it is the only material present, confirming the presence of the Vocal Delay layer (229.1 – 246.6 seconds) in the right channel required no filtering. Figure 47 illustrates the bounds of the right channels Vocal Delay material in sub-section 1 as well as a filtered version of the sensory dissonance graph for comparison to Figure 46. In sub-section 3, the presence of layering is quite obvious because it is part of Hiller's use of imitation with the reoccurrence of the Vocal Delay material. Figure 48 illustrates several timbre based feature profiles for sub-section 3 in both channels. The differences in timbre profiles for the left channel materials are clear, confirming the layer's presence from 269.9 to 310.4 seconds. The right channel trends are somewhat muddled by overlapping materials at both the beginning and end of the layer, yet 255.8 and 310.4 seconds still clearly demarcate points

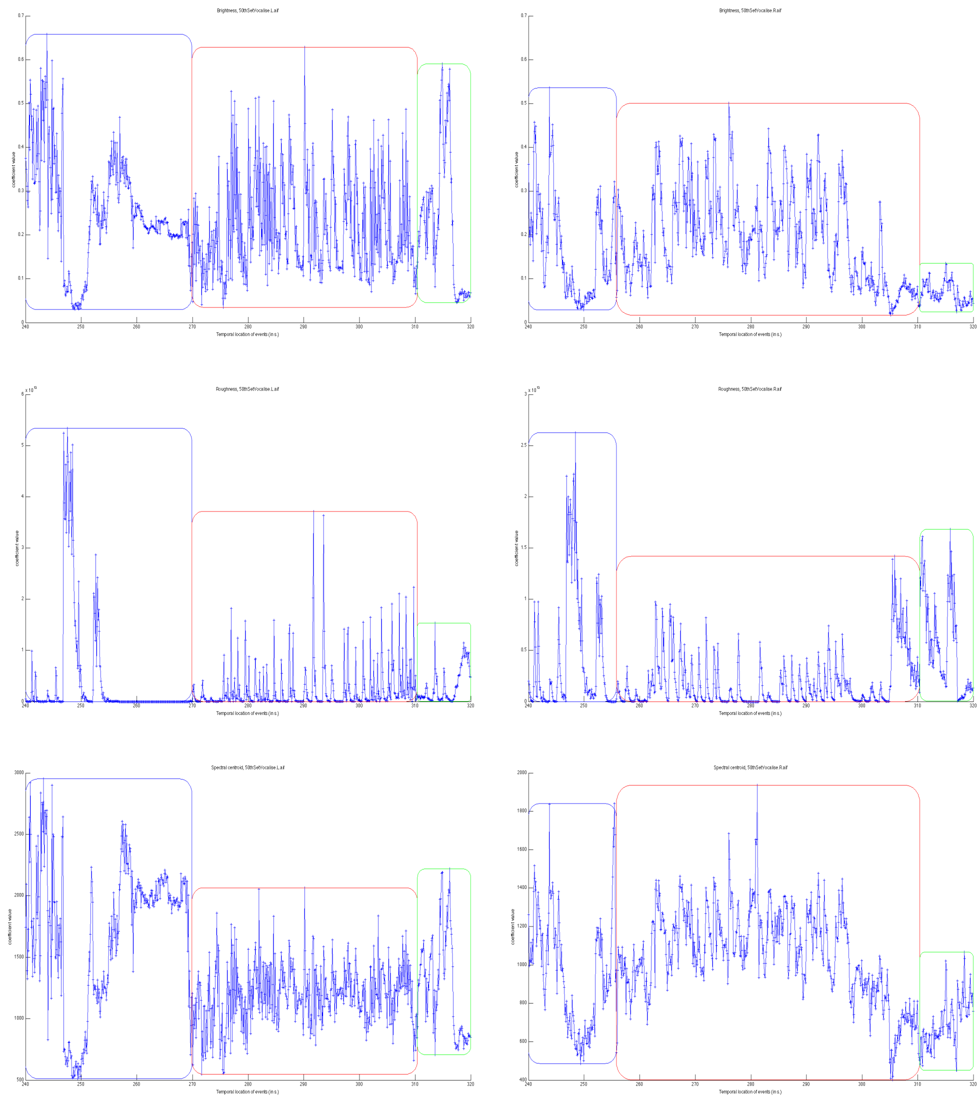


Figure 48 - Section 3 sub-section 3 [Top] Spectral Brightness, [Middle] Sensory Dissonance, and [Bottom] Spectral Centroid Graphs [Left = left channel, Right = right channel]

of change. The spectral centroid is perhaps the most useful for confirming the right channel Vocal Delay layer. Lastly, the re-emergence of the Vocal Synth Plus material in 285.2 seconds can be confirmed by the analysis results of Figure 49. Here, [Top] 285.2 seconds is identified as a point of significant spectral change, and [Bottom] 285.2 begins

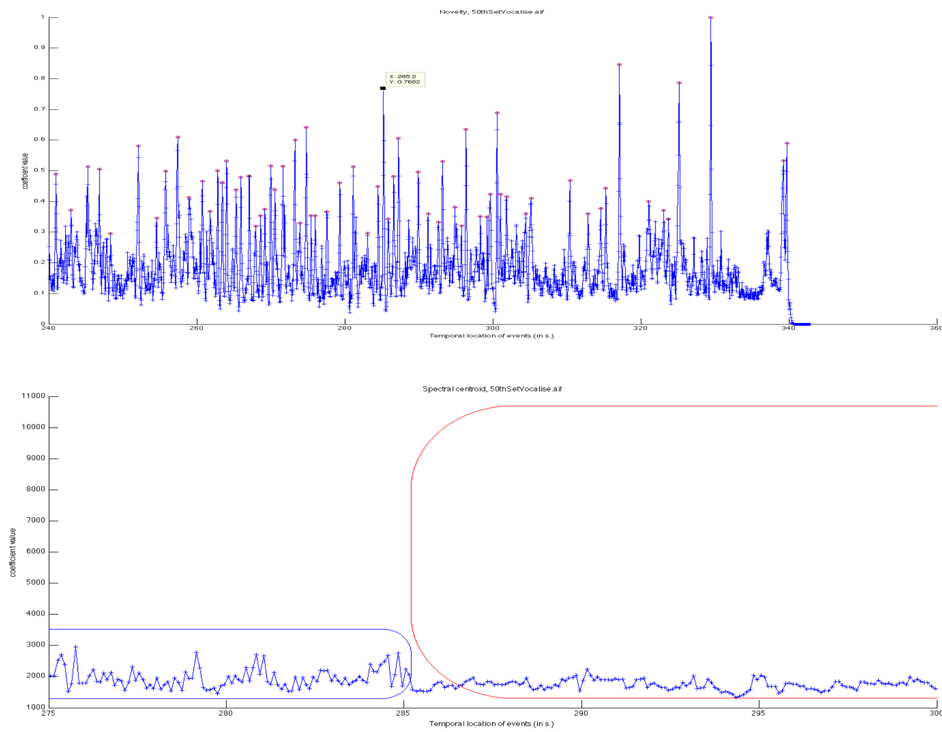


Figure 49 - Section3 sub-section 2 [Top] Spectral Novelty and [Bottom] filtered Spectral Centroid Graphs

a trend of less erratic spectral fluctuation indicative of the emergence of more steady state pitched materials.

4.2.4 – Event-level Analysis

Segmentation at event level is interpreted to be the identification of important events not heard as part of a stream/layer of materials or events that have sufficient separation/contrast with their surrounding materials to warrant separate identification. The segmentation of a work at the event level could encompass the segmentation of a work into all its component events, but there would be little gained and a great amount of effort wasted in such an approach. Since we hear clusters of similar events as a group in most cases, this approach seems to be a logical choice.

Since the majority of the work is composed of layers of similar events, and these layers are heard as such and not as separate occurrences of similar events, there are few events to make note of in this section of the analysis. In Section 1, sub-section 1 has two Noise Event Type A events that are presented with separation and contrast in terms of the spatialization approach taken. These events occur from 1.4 to 9.8 seconds and 11.75 to 21.1 seconds. Also in Section 1, in sub-section 4 (the last sub-section) the Recorded Noise Event A warrants mention for contrast and use throughout the work. This event occurs in both channels from 70.3 to 75 seconds. In Section 2, there is another occurrence of the Recorded Noise Event A, albeit with some modifications (reduced

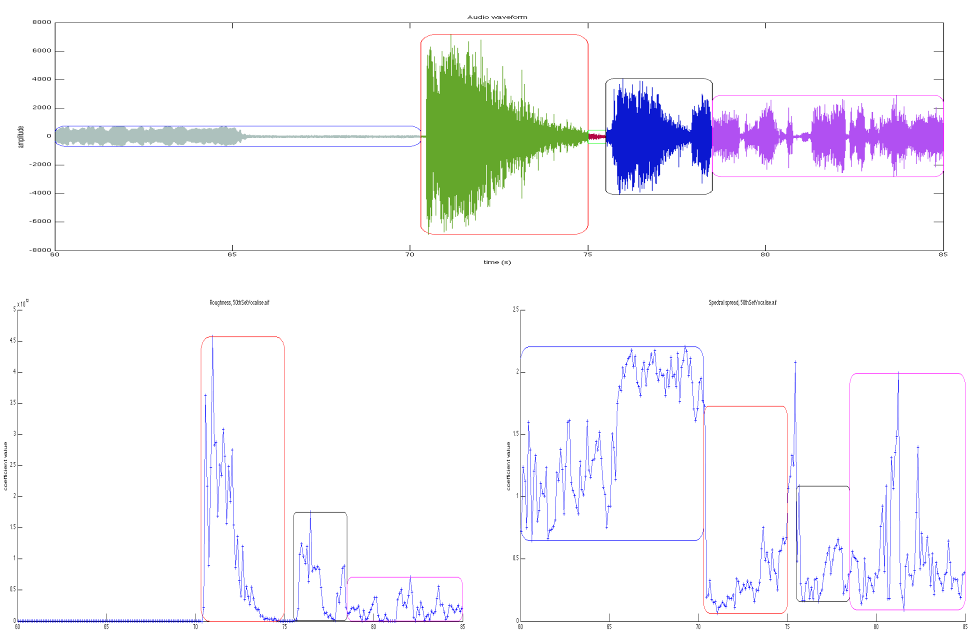


Figure 50 - 60 to 85 seconds filtered [Top] Segmented Waveform, [Bottom Left] Sensory Dissonance and [Bottom Right] Spectral Spread Graph

amplitude and at a higher frequency level), from 75.5 to 78.5 seconds. In sub-section 3 of Section 2 from 218.75 to 223.1 seconds, the Recorded Noise Event A occurs again in the left channel. This is followed by another *concréte* or recorded event, Recorded Noise Event B, in the left channel from 223.75 – 226.25 seconds. Finally, Section 3 sub-section 2 consists of two separate events: 246.8 – 252.1 seconds and 252.1 – 258.8 seconds.

Confirming either separation or contrast with most of these events via MIR was not difficult since the events were selected for the qualities of separation and contrast. Figure 20 and Figure 23 confirm the events present in Section 1 sub-section 1 to have both separation and contrast. The Recorded Noise Events at 70.3 to 75 seconds and 75.5 to 78.5 seconds are illustrated in Figure 50. Although the majority of the other materials present were able to be filtered out, only so much could be done to isolate the second event in the analysis due to the frequency range it occupies. Still, there are enough similarities between the events in the analysis result. The Recorded Noise Event A

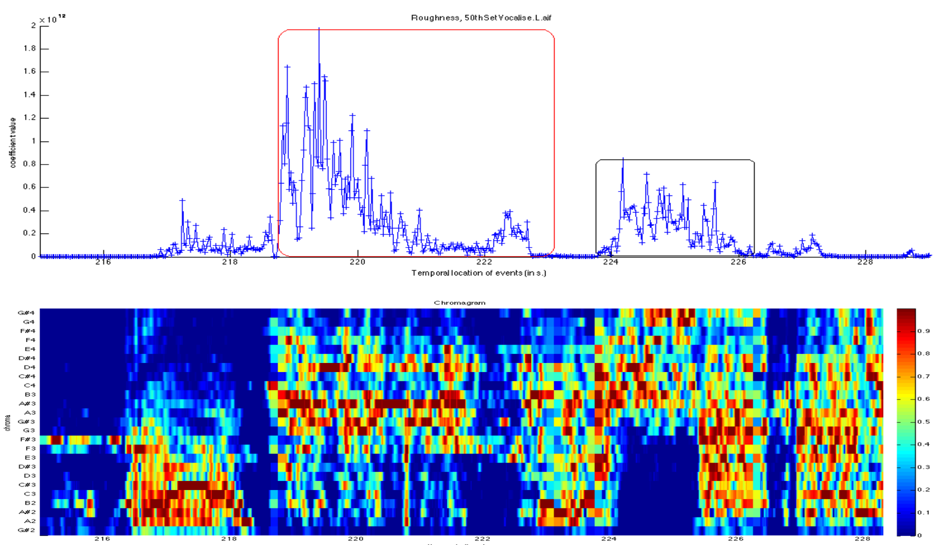


Figure 51 - Section 2 sub-section 3 filtered right channel [Top] Sensory Dissonance Graph and [Bottom] Chromagram

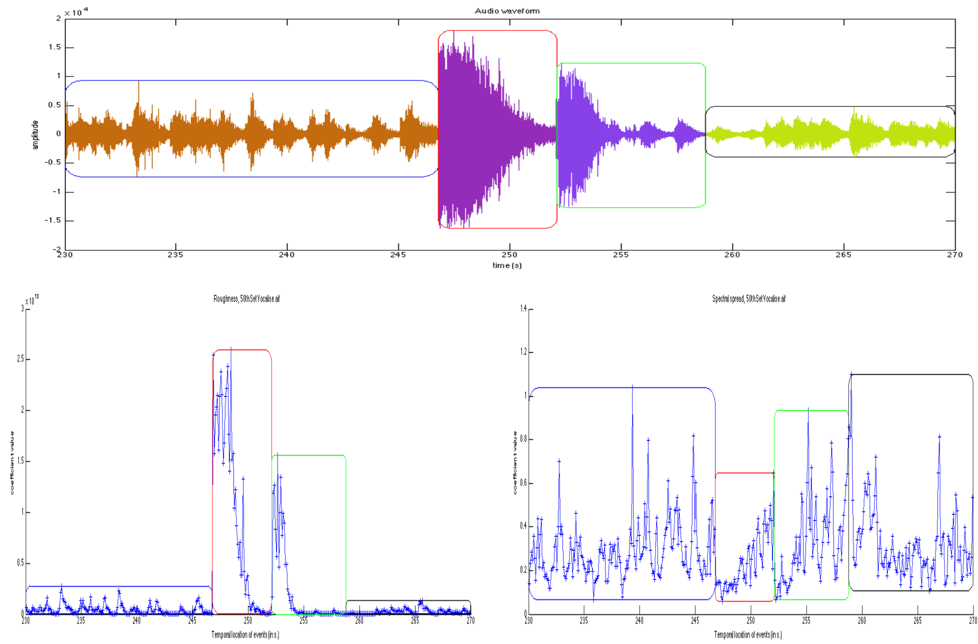


Figure 52 – 230 to 270 filtered [Top] Segmented Waveform, [Bottom Left] Sensory Dissonance, and [Bottom Right] Spectral Spread Graphs

(218.75 – 223.1 seconds) and Recorded Noise Event B event (223.75 – 226.25 seconds) are well represented by Figure 51, which illustrates differences in sensory dissonance as well as pitch content of the events and surrounding materials. The last two events addressed are a Recorded Noise Event A and Recorded Noise Event B, which are very similar in arrangement to the last two events addressed (218.75 – 223.1 and 223.75 – 226.25 seconds): they are presented in succession and the first event is at a higher amplitude and at a lower frequency level. Figure 52 illustrates the characteristic amplitude and dissonance profiles for these last Recorded Noise Event presentations at 246.8 – 252.1 and 252.1 – 258.8 seconds.

4.3 - Analytical Commentary

4.3.1 – Form (High to Mid-level)

As Hiller intimated in his program notes for *Vocalise*, the work is in three parts. These parts are labeled A, B, and C respectively, which would make the piece Through Composed in terms of form. Using letter names or descriptors of function where appropriate, the sub-sections of each section would be addressed as follows:

- Section 1 or A (0 – 75 seconds):
  - Introduction (0 – 17.7 seconds), *a* (17.7 – 45 seconds), *a'* (45 – 70.3 seconds), and close (70.3 – 75 seconds)
- Section 2 or B (75 – 229.1 seconds):
  - Transitional–*b* (75 – 93.47 seconds), *b* (93.47 – 216.2 seconds), and *c* (216.2 – 229.1 seconds)
- Section 3 or C (229.1 – 343 seconds):
  - Transitional–*d* (229.1 – 246.6 seconds), interruption (246.6 – 255.8 seconds), *d* (255.8 – 310.4 seconds), and coda (310.4 –343 seconds)

4.3.2 – Functions and Characteristics (Mid to Event-level)

At the Mid-level, the function and characteristics for each section have been partially addressed. The labels used in the previous section indicate much about the function of each section. Starting in Section A, the first sub-section fits with what we expect from introductory material: it is not too interesting, it is not in the foreground, and it fulfills the task of introducing the sonic pallet of the piece. Each of the noise-based events that make up the introduction are unique in that their approach to spatialization is not static – each of the events drift from the right to left channels over the course of their duration. Each event could be described to emerge from and release into silence except for the second event's release, which is obscured by the overlapping Vocal Synth material. The sub-sections *a* and *a*' are thematic in nature, presenting the listener with the Vocal Synth material as the theme, which is accompanied by the Noise Event Type B layer of materials. The theme is presented in the foreground with an economy of pitches and clear attacks for each note event. The theme consists of two voices in each of its occurrences with the voices positioned separately in each channel. The Noise Event Type B layer accompanies the theme in both channels in a rhythmically erratic and asynchronous fashion with short bursts of noise characterized by quick attack and decay times and no release. The last sub-section of section A is the close. Descriptors like interruption or derailment could be associated with this sub-section, but in the narrative of the work, this sub-section serves as a close to section A. Much like a close in sonata form, the close sub-section here brings finality to a section that could not have been achieved by using the theme as ending material. The close is made up of two materials: the Recorded Noise Event A, which is a sound used several times throughout the work; and a sonority presented with a quickly repeated ramping material. The Recorded Noise Event A appears to be a recorded percussive event – something either struck or dropped, with moderately quick attack and decay times, and moderate release time. The processing used on the event is likely tape speed variation, equalization, and reverb. The ramping sonority is presented with the Recorded Noise Event A in the foreground, and was likely constructed from a single reversed and speed varied recorded event. The timbre of the material has a ringing glass-like quality.

In section B, the first sub-section serves as a transition to the *b* sub-section. Some of the Section B's primary sonic materials emerge in this sub-section, however, there is too much activity, the section A materials are too present, and the sub-section too short to be called *b*. With its mix of section A (Noise Event Type B) and B (*Concréte* Materials/Loop, High Vocal Synth) materials, it prepares the listener for the following *b* sub-section. Although very different from the *a* sub-section in both type of material presented and how it is developed throughout its presentation, Sub-section b is thematic in nature. Although the *Concréte* Materials/Loop and High Vocal Synth materials hold primacy in the texture of event, they share in the label of theme for sub-section *b* with the idea of repetition and looping. The materials presented are developed by the ever-changing superimposition of elements with different lengths for each of the repeated or looped segments of material. This is done with a building trajectory, leading into sub-section *c,* by means of increasing amplitude and activity/complexity of the texture of events. Section B could be considered in Robert Frank's terminology to be built of a lattice of Non-Aligned/Repeating layers with some Non-Aligned/Non-Repeating materials added in toward the end of the sub-section to achieve the desired mass effect. Sub-section *c* is also thematic in nature with the *Concréte* Stutter material as the focus and materials from *b* and isolated events/gestures (Repetitive Attacking Bass gesture, Recorded Noise Event A, and Recorded Noise Event B) as accompaniment. Sub-section *c* is the climax of the piece referred to by Hiller in his program notes and has a trajectory of decreasing amplitude, activity, and spectral width over the course of its duration.

Section B is the most active section of the piece in terms of layers. Starting in sub-section 1 or transition-*b,* there are the materials that carry-over from section A: the Noise Event Type B material and the Recorded Noise Event A. The Noise Event Type B material here occurs in both channels in the mid-ground with similar event characteristics as when it appeared in section A, and with no contouring of the layer – it maintains the mid-ground position and amplitude for the entirety of its duration. The Recorded Noise Event A at the start of transition-*b* is modified in comparison to the Recorded Noise Event A presented in the close; here, aside from being presented only in the left channel, it simply appears to be at a higher pitch level. This is likely the result of a slightly different tape speed variation approach or setting. The section B materials that begin in the transition-*b* sub-section are the High Vocal Synth and *Concréte* Materials/Loop layers, both of which are presented in the left channel and last through the start of sub-section *c*, albeit with some breaks in the High Vocal Synth material. This mirrored and cyclic variation of the Vocal Synth material shares timbral makeup as well as some pitch material with the *a* theme. This layer occurs in the 3000 to 7500 Hz range and stays in a mid-ground position for its duration, mainly audible through its separation of tessitura from the other materials active in section B. The *Concréte* Materials/Loop layer is presented in the mid to background at a somewhat low amplitude when it first appears, and over its duration moves to a foreground position at a higher amplitude. This layer starts simply as the ~16.6 second loop of *concréte* material, but other recorded materials are added progressively as it approaches sub-section *c*. The loop is constructed of materials that could be described as either a door squeaking/creaking or events of a percussive nature that share a timbral makeup with the Recorded Noise Event A, but have been contoured and processed differently. To this, more recorded materials are added that are of a more piercing quality: percussive glass and perhaps metallic sounds.

Several layers begin at the start of section *b*: the Noise Event Type C layer, the repetitive ramping event layer, and the "bass line" layer. The Noise Event Type C layer consists of 14 Noise Event Type C events, presented in the mid to foreground distance in the left channel. The events are progressively reduced in amplitude over the duration of the layer with the repetitions averaging every ~5 seconds. The events have an almost instant attack with a moderately quick decay, no sustain, and a moderate release. The ramping event layer is the reverse of the Noise Event Type C layer in some ways: the envelope and spectral shape of events (hence the ramp label), and the amplitude contour for the layer, which starts low with a generally increasing trend over the duration of the layer. Otherwise, the ramping events themselves differ from the Noise Event Type C events in that they are more sonorous and not totally noise based. Additionally, they have linear based amplitude changes while the Noise Events layer is of a more organic exponential shape, likely due to the noise events being processed with a natural sounding reverb. Also positioned in the left channel at a distance of mid to foreground, the ramping events layer consists of 30 ramp events with the repetitions averaging every ~4 seconds. The "bass line" layer is unique in its similarity to the introduction of the piece –

it moves or pans through both channels and is not simply positioned. This synthesized mid to foreground layer could be described as gentle swells that drift from the left to the right channel, alternating from pitch class C to F.

In sub-section *c*, as mentioned, the *Concréte* Based Stuttering material presented in the left channel is the theme. It consists of a re-attacking (stutter) percussive recording that is processed via filtering to achieve a descending glissandi effect in multiple streams. The recording(s) appear to have similar timbral makeup to some of the percussive elements of the *Concréte* Materials/Loop. This layer remains in the foreground for the duration of sub-section *c*, but moves to the background over the duration of the transition-*d* sub-section. The spectral width of the layer decreases rapidly from the start of *c* to 227.5 seconds and then abruptly shifts back to a wider state before beginning a general decrease that ends with the layer. The supporting materials of sub-section *c* not already elaborated upon are all relatively short and unique in the local context of sub-section *c.* The repetitive attacking bass gesture occurs abruptly below 400 Hz and was likely constructed via processing by extended tape speed variation some of the filtered materials from the *Concréte* Based Stutter layer's construction. Immediately following this event is another Recorded Noise Event A, which appears to be identical to the event that occurred in the close. Lastly, the Recorded Noise Event B event occurs immediately following the last ramp event in sub-section *c.* This event has qualities similar to that of the Recorded Noise Event A, yet it clearly has a different, gentler attack and what seems to be a slightly different timbral makeup. It is difficult to be sure, concerning the timbre, considering the amount of activity in that frequency range at the time of the event.

Section C begins like section B, with a transitional section made up of materials from the previous and new sections. Transition-*d*, like transition-*b,* prepares the listener for the following thematic sub-section. Aside from the carry-over material from section B, this is the first occurrence of the Vocal Delay material, which occurs in the mid to foreground in both channels. The material in each channel is different, however, they appear to be the same material, simply processed differently – the left channel is composed of short vocal utterances processed with delay, and the right channel appears to be the left channel material further processed with reverb and equalization. The vocal utterances used in the Vocal Delay material vary and due to the processing, it is difficult to determine if a particular text is being conveyed. All of the utterances have a breathy whispered quality, appear to be recordings of a male speaker, and sound as if recorded with a close microphone position. The following sub-section, interruption, functions as an abrupt sidebar for the listener, and is long enough to create some separation between the materials to come and the recent climax of the piece. The interruption consists of two events, a Recorded Noise Event A and a Recorded Noise Event B, that are ~5.25 and ~5 seconds long respectively. This combination of events appeared in sub-section *c*, but with lack of masking in the interruption sub-section, much more can be observed about the Recorded Noise Event B. The event consists of more than just a single percussive event as in the Recorded Noise Event A; while still percussive in nature, it is composed of several attacks and seems likely to be the product of a dropped object cascading off other objects/surfaces before coming to rest. The events of the interruption sub-section could be described to function as an interruption for the Recorded Noise Event A and as a reiteration and affirmation for the Recorded Noise Event B event. The *d* sub-section is thematic in nature, presenting the foreground Vocal Delay material in imitation without supporting materials. The Vocal Delay material is once again separated into different channels with different levels of processing. The more heavily processed right channel enters first, at the beginning of the section, with the less processed version entering in the left channel ~14.5 seconds later. The coda, while not a prolongation of tonic as in the common practice, is a prolongation of the primary compositional and timbral idea of the work. The Vocal Synth Plus material emerges gradually from the background in sub-section *d* to the foreground by the start of the coda. The number of voices here is difficult to determine, however, the focus is not on singular voices, but on the homophonic presentation of several sonorities to close the work. Low frequency noise elements accompany the Vocal Synth material briefly at the end of sub-section *d* and start of the coda, as well as end the piece after the final Vocal Synth sonority.

4.4 – Salient Features

4.4.1 – Materials

As the title would suggest, vocal and vocal-like materials are central to Hiller's *Vocalise.* Materials described as vocal or vocal-like are primary materials in two of the three sections of the work, and share primacy in the 3rd section. The *a* theme, while synthesized, is constructed with an acoustic model within the range and capabilities of an actual vocalist. The imposed human boundaries are disregarded in the development of the *a* material in the B section, taking into consideration the lack of such limitations in electro-acoustic music. In section C, the Vocal Delay materials continue the vocal "theme."

Vocal or vocal-like material is the focus, and the majority of the piece can be described as interactions of vocal and noise based elements. The noise based elements mostly function in a supporting capacity, either directly as an accompanying material or contributing to interruptions and impacts: Noise Event Type B materials, Recorded Noise Event A, Noise Event Type C materials, and Recorded Noise Event B events. Considering the number of appearances the more important noise based materials are the Recorded Noise Event A and Noise Event Type B material.

4.4.2 – Processing Methods

The following is a record of the some of the identifiable processes that were utilized in the construction and transformation of materials for *Vocalise.* Listed with each processing type are some materials that are the result of that processing.

- Tape speed variation => Recorded Noise Event A and Recorded Noise Event B
- Reverb => Recorded Noise Event A, Recorded Noise Event B, and Noise Event Type C
- Equalization => *Concréte* Materials/Loop and *Concréte* Based Stuttering
- Delay => Vocal Delay
- Filtering => *Concréte* Based Stuttering and Re-attacking Bass Gesture

4.4.3 – Compositional Methods

This section makes note of some compositional methods or techniques utilized in the construction of *Vocalise.* The areas of interest are: the construction of the *Concréte* Materials/Loop loop, the counterpoint between the Noise Event Type C and ramping event layer, the construction of the *Concréte* Stuttering materials glissandi, and the use of semi-dry and wet signals in the transition-*d* and *d* sub-sections.

The construction of a loop is akin to the composition of a set, row, or minimalist's motivic materials; it must be done with care. The composer must take into consideration all possible ways the material will be manipulated or combined with other materials, so as to maximize the effectiveness and minimize the pitfalls of using such economical source material constructions. Seasoned listeners can identify a weak set, row, motive, or loop. Without making judgments of Hiller's loop of *concréte* materials, listed below are some aspects of the loop's construction that were a benefit to the prolonged presence of looped materials.

- The loop was constructed mostly of materials with similar characteristics (timbre, event shape, and tessitura), presented arrhythmically.
- The few events that did not fit into the above description (the door squeak/creak events) were presented with ample space in between similarly contrasting events.
- The loop has a faux repeat or restart point in the middle of the loop. The door squeak/creak that starts the loop is used again at the mid-point of the loop causing the listener to be confused about when the loop actually repeats.

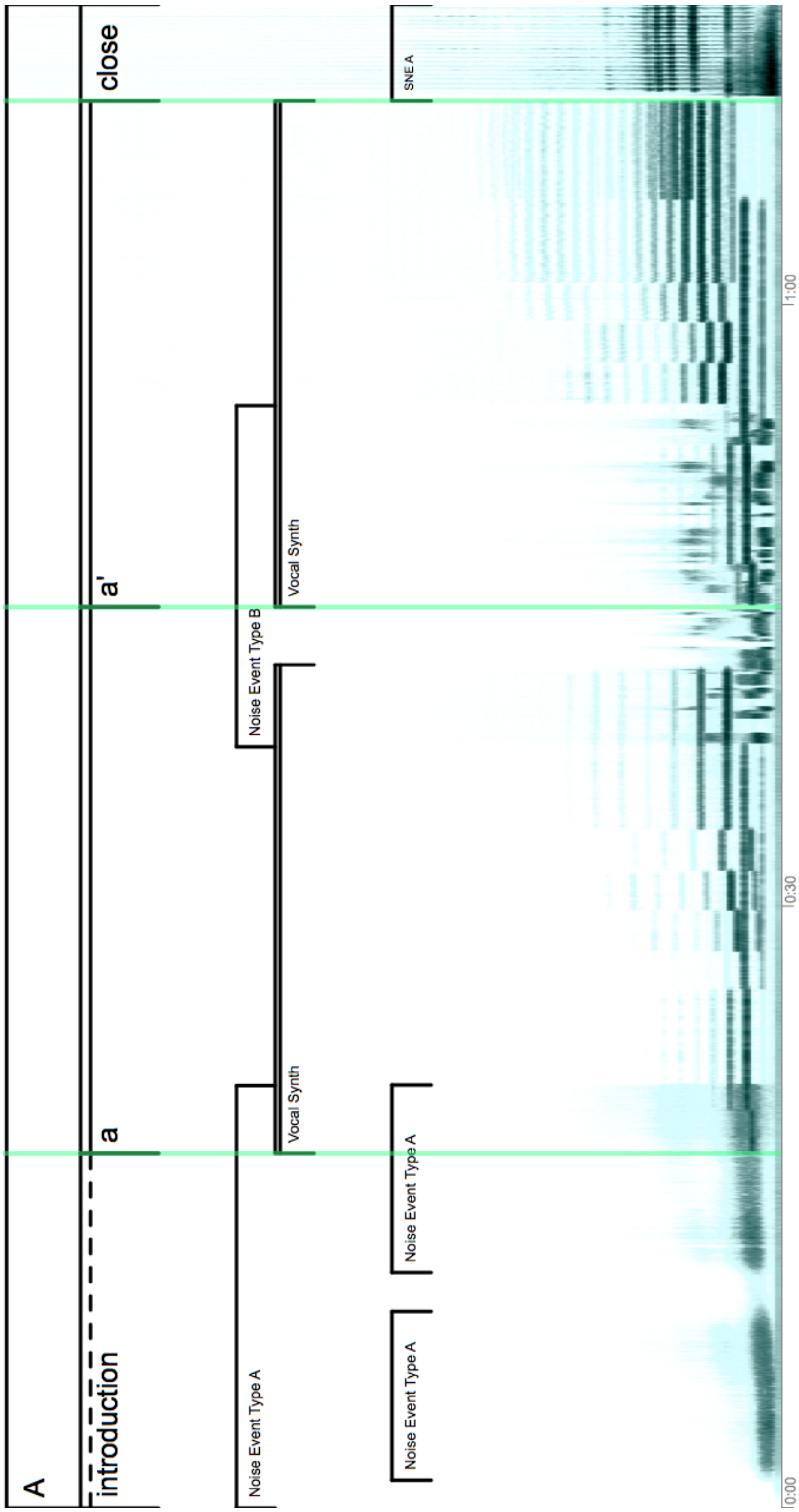
Hiller also minimizes the possibilities of any weaknesses becoming apparent by having other materials active during the loops presentation.

There is an interesting symmetry in the counterpoint between the Noise Event Type C and ramping event layers of section B. The Noise Event Type C layer has a sharp attack, lengthy release, and has a similar layer contour of starting at a high amplitude and ending with a low amplitude, and the ramping event layer has a lengthy attack slope, sharp release, and a likewise similar layer contour of low to high amplitudes. This scenario has the potential to be interesting or quite rigid depending on how it is put to use. Hiller does a number of things to maximize the effectiveness of this economic arrangement of materials: the layers are not synchronized, the layers repeat at different increments with a high common denominator, the layers are different lengths, and the layers do not start or stop together.

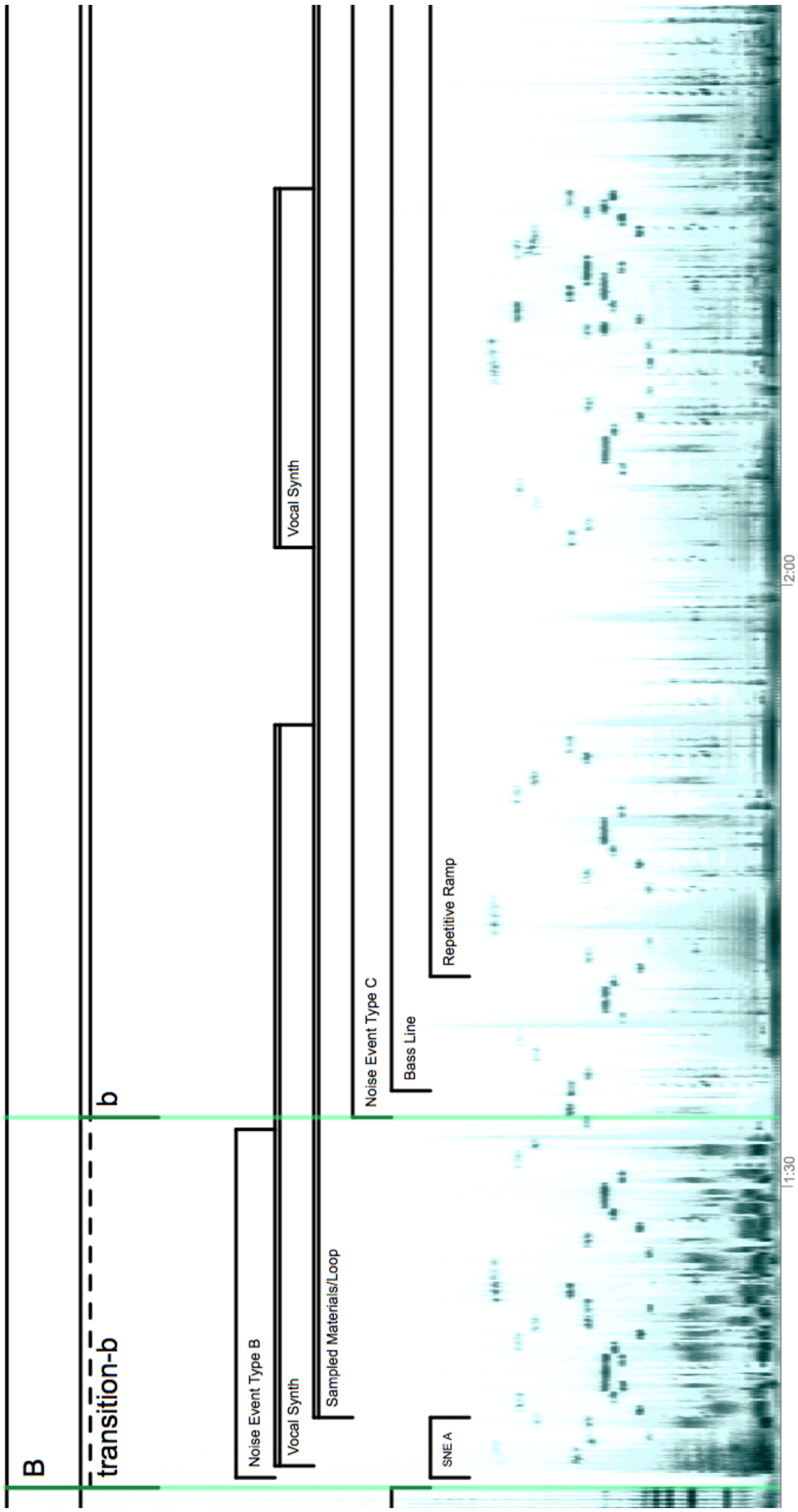
The *Concréte* Stuttering theme of sub-section *c* has an interesting glissandi aspect of its construction that was well executed. Considering the equipment likely available at the time, for this material to be constructed either each glissandi stream would have to be recorded separately and mixed together with the source material to create the final result, or each glissandi stream was mixed into the source recording before the creation of the next glissandi stream to achieve the final result. The later option seems more likely, or perhaps in practice, it was a mix of the two approaches – doing a mix after creating a group of glissandi streams. In either case, this material required several layers of processing and/or mixing to create, and the final result was an effective yet non-obtrusive multi-stream effect that added depth to the layer.

Lastly, the arrangements of semi-dry and wet signals in the transition-*d* and *d* sub-sections are worthy of note. In these sub-sections, the vocal utterances source material appears to be processed with delay to create one layer of material (semi-dry), and then this material is processed further with reverb and equalization to form another layer of material (wet.) In the transition-*d* sub-section, these layers of material are presented simultaneously, and quite exposed with the semi-dry signal in the left channel and the wet signal in the right channel. Here the listener can perceive which layer is the less processed version and which sounds less like what the original must have sounded like. The transition-*d* sub-section prepares the listener for the similar arrangement of materials in sub-section *d.* It is interesting that Hiller chooses to begin the imitative passage of sub-section *d* with the more processed version of the Vocal Delay material, having the less processed version serve as the imitating voice.

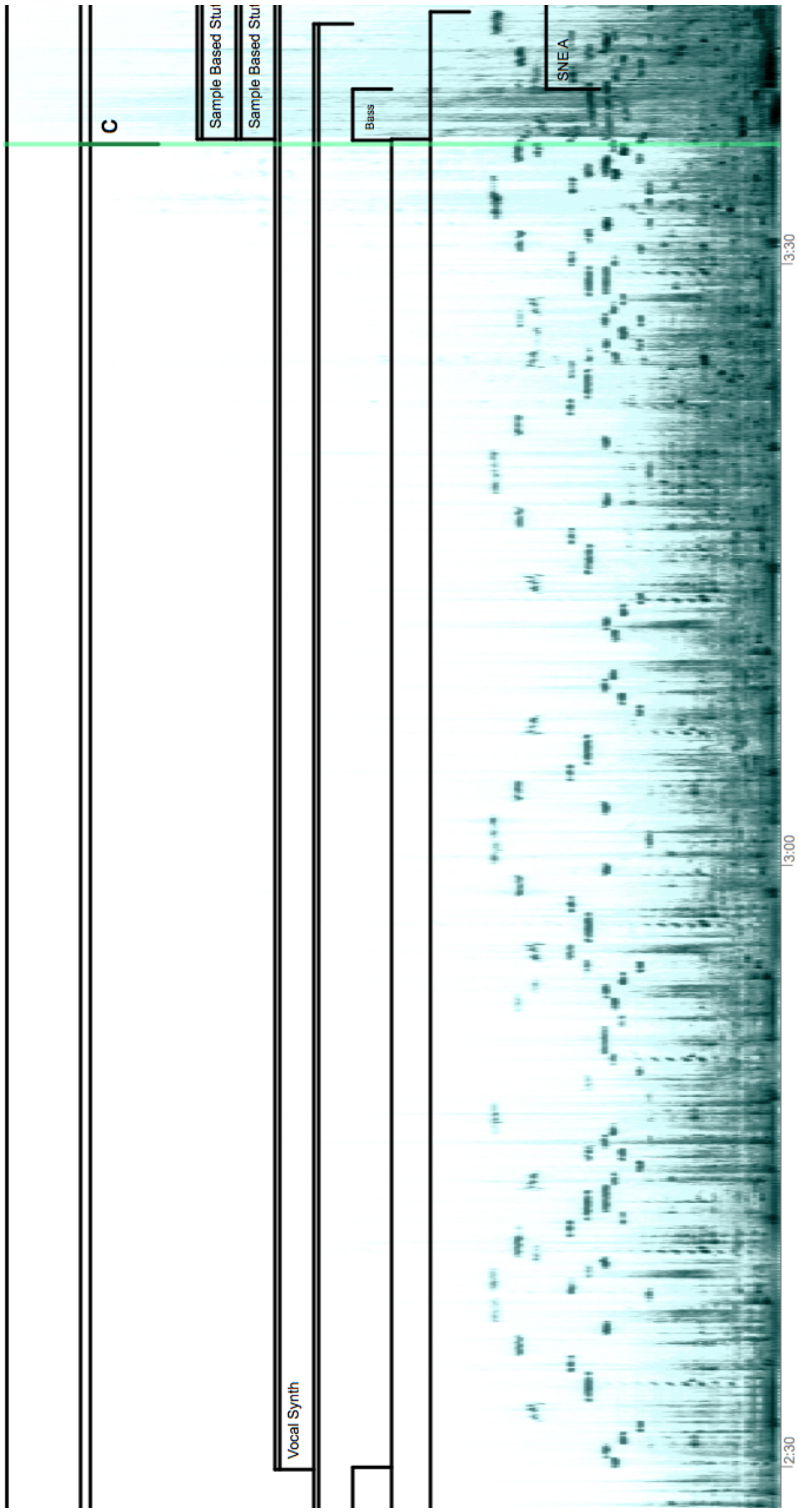
4.5 - Analysis Score



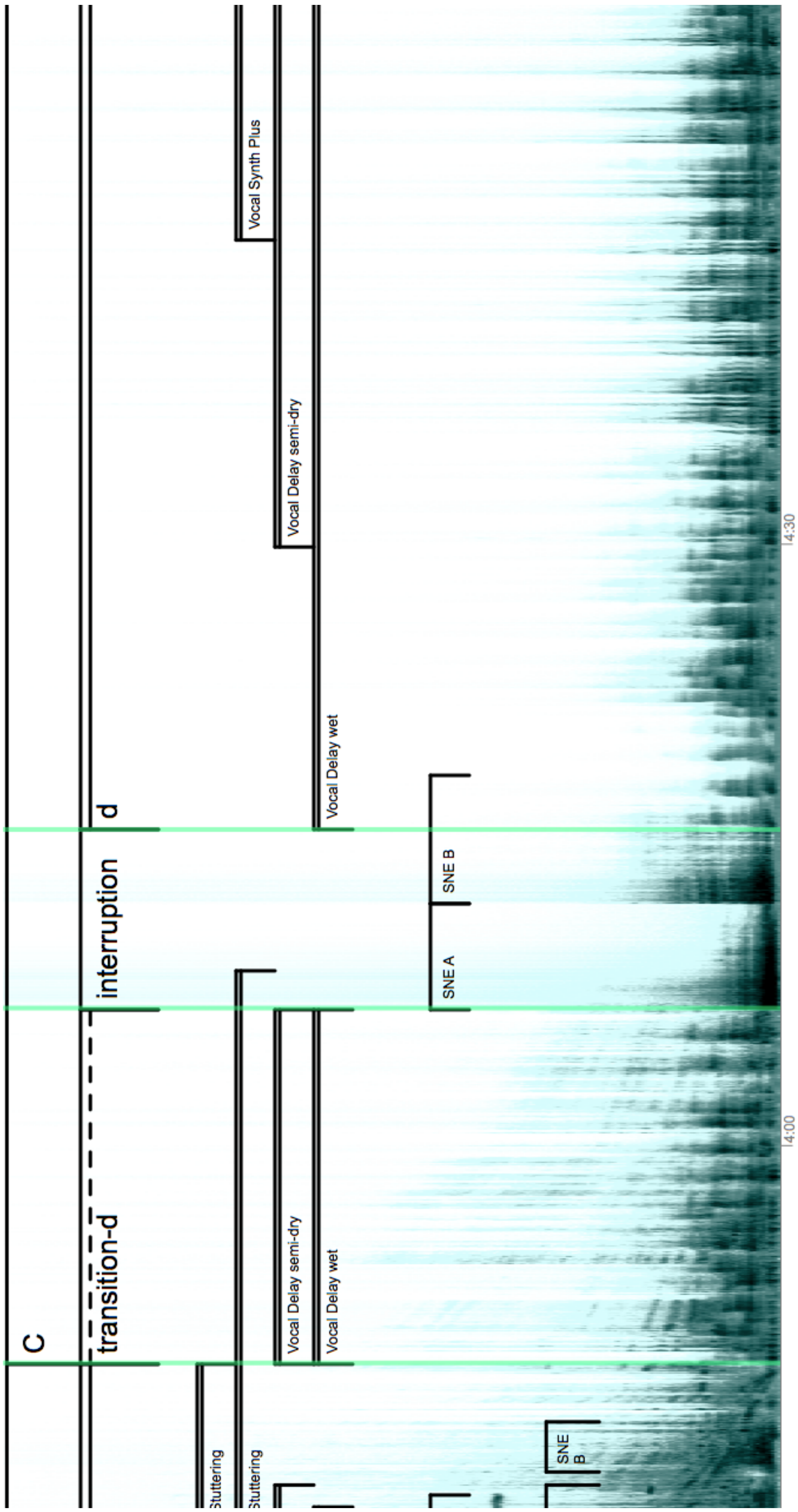
Analysis Score (Cont.)



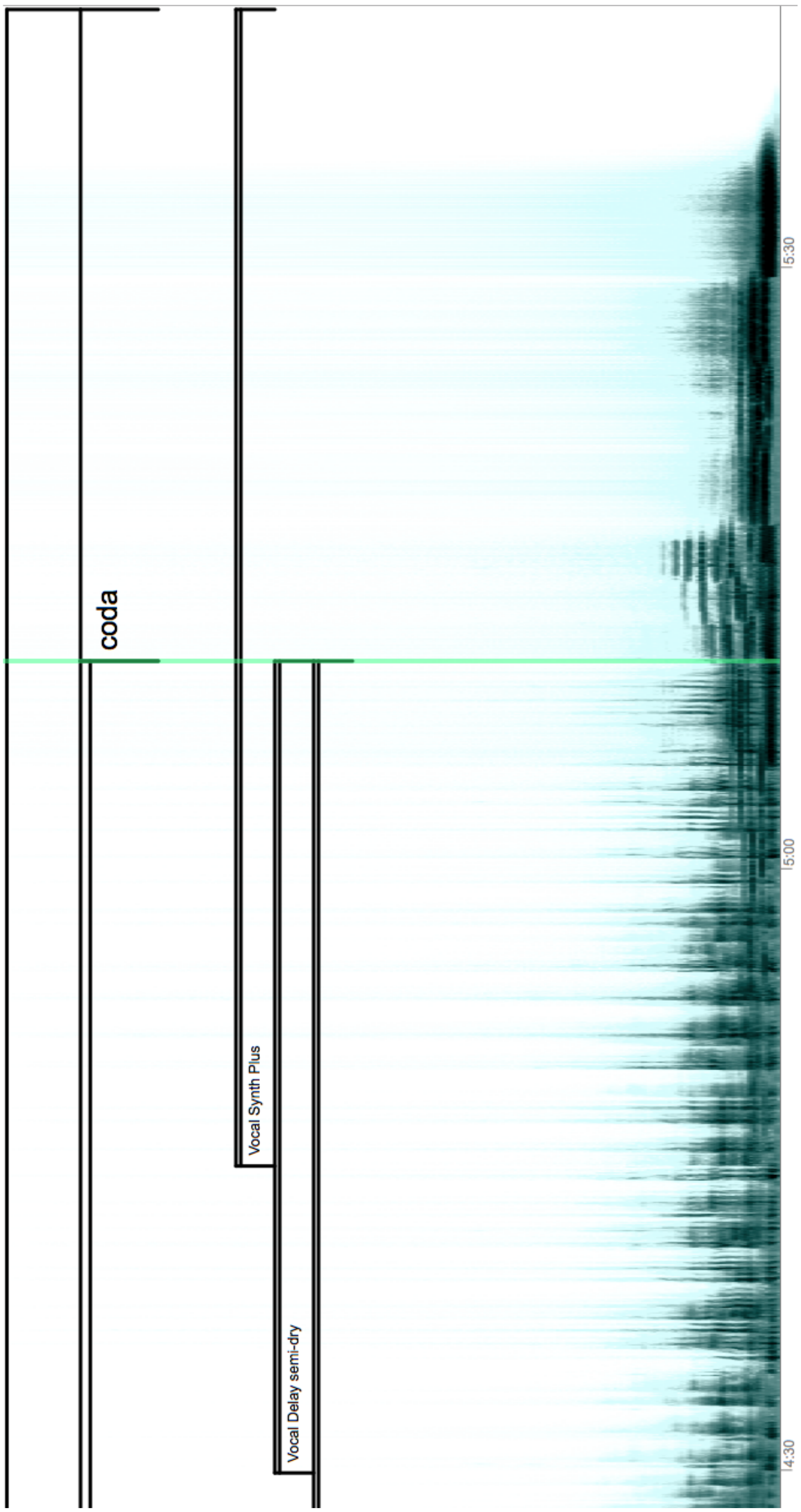
Analysis Score (Cont.)



Analysis Score (Cont.)



Analysis Score (Cont.)



4.6 – Concluding Statements

The analysis method used was well matched for the examination of Hiller's *Vocalise.* Segmenting the work into layers as well as formal divisions, was particularly useful for understanding the way the work was constructed. SQEMA provided a logical framework for how to approach the analysis, the listening-based analysis methodologies provided examples of appropriate descriptive language, and MIR analysis provided useful confirmation and empirical insight into the qualities of the materials analyzed.

Although the approach used in the study was based on SQEMA (a system where a listening analysis could be used to confirm/evaluate an MIR analysis), it is important to reiterate that in this study, Hiller's *Vocalise* was analyzed through listening and used MIR as confirmation. For this study, the use of MIR was not necessary to determine divisions and layers of the work. The MIR analysis confirmed all aurally identified divisions and layers, yet no divisions or layers were discovered with MIR that were not already identified via listening. MIR was an invaluable tool for understanding how the qualities of materials related. Having a measurable comparison of musical parameters and their evolution was extremely useful. It was important as a learning experience to manually examine and filter the results of the MIR analyses, however, there are some aspects of working with MIR that could be automated: identification of trends, clustering of materials according to trends, and feature comparisons. Although research into the complete automated analysis of music via MIR is being pursued, such endeavors are not a replacement for an adept music analyst.

Based on the work here and on some experimentation, there are some MIR tools that would be useful for electro-acoustic music analysis that are not yet offered by the MIR analysis paradigm: a noise to pitched sound graph, similarity graph for a single frame to whole file comparison, and a similarity matrix between two files. A noise to pitched sound graph or measurement of how pitched a given spectrum is would have been useful in the analysis of *Vocalise*; it would have been especially useful in the analysis of section A, where the chromagram was examined to determine pitch saturation. A similarity graph that could test a single moment against a whole file or series of files would be very useful in many types of analysis. With such a tool, an analyst could easily test for occurrences of any sonic event as long as they have a model to use for the test. The similarity matrix is useful in the context of electro-acoustic music analysis, but a similarity matrix where the analyst could specify the files or file segments used would allow far greater control over the analysis being performed.

There is much research that could be based on the work done in this analysis: listening pedagogy (music appreciation and outreach), composing with MIR feedback, advanced interactivity in live electronics, and possibly development of the above mentioned MIR tools for use in electro-acoustic music analysis.

- Very few listeners have the intuitive capacity to grasp, while listening, that the materials presented in electro-acoustic music (or any modern musical dialect) are different from those composed in the common practice and should not be listened to with the same approach. The approach refers to what the listener listens for and how they expect the musical discourse to unfold. Many talented musicians have trouble with this as well; they listen for a tune and a beat where they will not find one, and are disgruntled. Addressing how to listen and what to listen for in electro-acoustic/modern music is the answer. Doing so in the music appreciation context makes sense, but also in outreach and preconcert lectures would make an impact.

- Hiller's *Vocalise* and many other works have commonalities in their MIR trends. These trends, of course, correlate to the qualities of the musical discourse and as such can be utilized in music composition. Having an informed awareness of qualities like spectral centroid, sensory dissonance, and metric stability, among others, could be useful in the composition process. After understanding the range and limits of a specific quality, composing contours/trajectories as a pre-compositional step and using those models in the composition of work could be quite fruitful. Tools could be created to analyze midi and audio data in real time to monitor and give feedback a composer goes about their work. There are also numerous possibilities for algorithmic composition informed by MIR analysis and algorithms driven by models of behavior.
- MIR analysis provides the capability of sophisticated interactivity in a live and studio context. This can include the triggering of events to take place when *X* quality reaches a certain state, when several qualities are in specific states, and score following. The data retrieved from MIR analysis can be used endless ways to actively manipulate events being generated.

The original goals for choosing this research topic have been met (understanding the problems and methods of electro-acoustic music analysis and informing algorithmic composition and electro-acoustic music pedagogy). Although a theoretical understanding of the issues associated with electro-acoustic music analysis is beneficial, it does not compare to the practical understanding of addressing those issues in an analysis. This research has impacted my own work in composition (electro-acoustic and algorithmic). My work has benefited from Hiller's examples of juxtaposition, his use of materials at various stages of processing, as well as his models of types of layered materials.

This investigation of electro-acoustic music analysis highlights the difficulties of analysis, explains some of the prevailing analysis methodologies, and demonstrates the use of a hybrid analysis methodology through an analytical case study. These difficulties and analysis methodologies correlate to electro-acoustic music in general as well as the piece chosen for the analytical case study, yet there is much work left to do in the study of electro-acoustic music. As electro-acoustic music is a broad genre, there is great potential for sub-genre specific issues and specialized approaches to analysis. Mixed medium (acoustic instrument and/or video with electro-acoustic music) and interactive works in particular would require further investigation into their sub-genre specific difficulties and the best possible analysis methodologies. These wide-ranging possibilities for electro-acoustic music as a genre dictate that its study is perhaps best approached with the general principle of evaluating each work without assuming knowledge based on its sub-genre, aesthetic, or methods of its composition – exactly how one would approach the analysis of an acoustic work.

Bibliography

Atkinson, Simon. "Interpretation and Musical Signification in Acousmatic Listening." *Organised Sound* 12.02 (2007): 113–122. Web.

Blackburn, M. "*Composing From Spectromorphological Vocabulary: Proposed Application, Pedagogy and Metadata*." *Unpublished paper*. Print.

Bohn, J. M. "An Overview of the Music of Lejaren Hiller and an Examination of His Early Works Involving Technology." 276. Print.

Bossis, B. "The Analysis of Electro-acoustic Music: From Sources to Invariants." *Organised Sound* 11.02 (2006): 101–112. Web.

Boura, V. "The Rhetorical Method for the Critical Appraisal of Electro-acoustic Structures." *Journal of New Music Research* 36.2 (2007): 115–138. Web.

Camilleri, L., and D. Smalley. "The Analysis of Electro-acoustic Music: Introduction." *Journal of New Music Research* 27.1-2 (1998): 3–12. Web.

Chadabe, J. *Electric Sound*. Pearson College Division, 1997. Print.

Couprie, P. "(Re) Presenting Electro-acoustic Music." *Organised Sound* 11.02 (2006): 119–123. Web.

Couprie, P. "Graphical Representation: an Analytical and Publication Tool for Electro-acoustic Music." *Organised Sound* 9.01 (2004): 109–113. Web.

Couprie, P. "Three Analysis Models for L'Oiseau Moqueur , One of the Trois Rêves D'Oiseau François Bayle." *Organised Sound* 4.1 (1999): 3–14. Web.

Couprie, P. "Three Analysis Models for L'Oiseau Moqueur, One of the Trois Rêves D'Oiseau François Bayle." *Organised Sound* 4.01 (1999): 3–14. Web.

Emmerson, S. T. "Analysis and the Composition of Electro-Acoustic Music." (1982). Print.

Franks, R. "Temporal Elements: a Cognitive System of Analysis for Electro-Acoustic Music." *ICMC 2000* (2000). Print.

Gang, R. et al. "A Real-Time Signal Processing Framework of Musical Expressive Feature Extraction Using MATLAB." *ISMIR* (2011). Print.

Gayou, É. "Analysing and Transcribing Electro-acoustic Music: the Experience of

the *Portraits Polychromes* Of GRM.” *Organised Sound* 11.2 (2006): 125–129. Print.
Godøy, R. “Chunking Sound for Musical Analysis.” *Unpublished paper*. Print.

Godøy, R. “Gestural-Sonorous Objects: Embodied Extensions of Schaeffer’s Conceptual Apparatus.” *Organised Sound* 11.2 (2006): 149–157. Print.

Harchanko, J. “Spectro-Morphology and Structure. an Analysis of Gilles Gobeil's *Le Vertige Inconnu*.” *eContact* (2003). Web.

Helmuth, M. “*‘Meeting the Free Dreamer’: an Exploration of Computer Music Composition*.” Print.

Helmuth, M. “Multidimensional Representation of Electro-acoustic Music.” *Journal of New Music Research* (1996). Print.

Hiller, L. “Electronic Music at the University of Illinois.” *Journal of Music Theory* 7.1 (1963). Print.

Hirst, D. “A Cognitive Framework for the Analysis of Acousmatic Music: Analysing Wind Chimes by Denis Smalley.” (2008). Print.

Hirst, D. “An Analytical Methodology for Acousmatic Music.” *ISMIR* (2004). Print.

Hirst, D. “*Developing Analysis Criteria Based on Denis Smalley’s Timbre Theories*.” *Unpublished paper*. Print.

Hirst, D. “The Use of MQ Plots in the Analysis of Electro-Acoustic Music.” *Unpublished paper* (2000). Print.

In Celebration of the 50th Anniversary of the University of Illinois Experimental Music Studios (1958-2008). *In Celebration of the 50th Anniversary of the University of Illinois Experimental Music Studios (1958-2008)*. School of Music, University of Illinois at Urbana-Champaign. Audio Recording.

Landy, L. *Understanding the Art of Sound Organization*. 2007. Print.

Lartillot, O., and P. Toiviainen. “A Matlab Toolbox for Musical Feature Extraction From Audio.” *International Conference on Digital Audio Effects* (2007). Print.

Lerch, A. *An Introduction to Audio Content Analysis*. John Wiley & Sons, 2012. Print.

Li, Tao, M. Ogihara, and G. Tzanetakis. *Music Data Mining*. CRC Press, 2011. Print.

Licata, T. *Electro-acoustic Music*. Greenwood Press, 2002. Print.

Meyer, L. B. *Explaining Music*. Univ of California Press, 1973. Print.

Mountain, R. “Theories Market: Open for Trading.” *Organised Sound* 9.01 (2004): 15–26. Web.

Park, T. H., D. Hyman, P. Leonard, and P. Hermans. “Towards Comprehensive Framework for Electro-Acoustic Music Analysis.” *ICMC* (2011). Print.

Park, T. H., D. Hyman, P. Leonard, and W. Wu. “Systematic and Quantative Electro-Acoustic Music Analysis (Sqema).” (2010). Print.

Park, T. H., Z. Li, and W. Wu. “Easy Does It: the Electro-Acoustic Music Analysis Toolbox.” *ISMIR* (2009). Print.

Paulus, J., M. Müller, and A. Klapuri. “Audio-Based Music Structure Analysis.” *ISMIR* (2010). Print.

Reed, T. “I. Planes of Discourse in Fixed Media Electro-acoustic Music: a Comparative Study and Application of Analytical Approaches and II. Three Movements for String Orchestra.” 116. Print.

Roy, S. “Form and Referential Citation in a Work by Francis Dhomont.” *Organised Sound* 1.01 (1996): 29–41. Print.

Roy, S. “Functional and Implicative Analysis of Ombres Blanches*.” *Journal of New Music Research* (1998). Print.

Schafer, R. M. *The Tuning of the World*. 1977. Print.

Simoni, M. H., N. Adams, L. Landy et al. *Analytical Methods of Electro-acoustic Music*. 2006. Print.

Smalley, D. “Defining Timbre—Refining Timbre.” *Contemporary Music Review* (1994). Print.

Smalley, D. “Space-Form and the Acousmatic Image.” *Organised Sound* (2007). Print.

Smalley, D. “Spectro-Morphology and Structuring Processes.” *The language of electro-acoustic music* (1986). Print.

Smalley, D. “Spectromorphology: Explaining Sound-Shapes.” *Organised Sound* (1997). Print.

Smalley, Denis. “Space-Form and the Acousmatic Image.” *Organised Sound* 12.01 (2007): 35–58. Web.

Thoresen, L. “Auditive Analysis of Musical Structures.” 1985. Print.

Thoresen, L. “Sound-Objects, Values and Characters in Åke Parmerud’s *Les Objets Obscurs*, 3rd Section.” *Organised Sound* 14 (2009): 310–320. Print.

Thoresen, L. “Spectromorphological Analysis of Sound Objects: an Adaptation of Pierre Schaeffer’s Typomorphology.” *Organised Sound* 12.2 (2007): 129–141. Print.

Tzanetakis, G., and P. Cook. “MARSYAS: a Framework for Audio Analysis.” *Unpublished paper*. Print.

Wishart, T. *Audible Design*. Orpheus the Pantomime Limited, 1994. Print.

Wishart, T. *On Sonic Art*. Psychology Press, 1996. Print.

Wyatt, S. “Investigative Studies on Sound Diffusion/Projection.” *Journal SEAMUS* (2000)

Young, J. “Reflections on Sound Image Design in Electro-acoustic Music.” *Organised Sound* 12.1 (2007): 25–33. Print.

Young, J. “*Sound in Structure: Applying Spectromorphological Concepts*.” *Unpublished paper*. Print.

Young, J. “Sound Morphology and the Articulation of Structure in Electro-Acoustic Music.” *Organised Sound* 9.1 (2004): 7–14. Print.

Zattra, L. “The Identity of the Work: Agents and Processes of Electro-acoustic Music.” *Organised Sound* 11.2 (2006): 113–118. Print.

Appendix A – MIRtoolbox (version 1.3.4) / MATLAB® code for Figures

Figure 10 [Top]

```
a = miraudio('50thSetVocalise.aif', 'Mono', 2)
```

Figure 12

```
mf = mirmfcc('50thSetVocalise.aif', 'Frame', .1,'s',.1,'s','Bands', 20, 'Rank', 1:10)
```

Figure 13

```
aTest = miraudio('50thSetVocalise.aif', 'Extract', 55, 95);
s = mirsegment(aTest, 'MFCC', 'Rank',1:5)
```

Figure 14 [Top]

```
aTest = miraudio('50thSetVocalise.aif', 'Extract', 55, 95);
sc = mircentroid(mirsegment(aTest, 'MFCC', 'Rank',1:5), 'Frame',
        .1,'s',.1,'s')
```

[Middle]

```
sn = mirzerocross(mirsegment(aTest, 'MFCC', 'Rank',1:5), 'Frame')
```

[Bottom]

```
sb = mirbrightness(mirsegment(aTest, 'MFCC', 'Rank',1:5),
        'Frame', .1,'s',.1,'s')
```

Figure 15

```
aTest = miraudio('50thSetVocalise.aif', 'Extract', 200, 260);
s = mirsegment(aTest, 'MFCC', 'Rank',1:5)
```

Figure 16 [Top]

```
aTest = miraudio('50thSetVocalise.aif', 'Extract', 200, 260)
sc = mircentroid(mirsegment(aTest, 'MFCC', 'Rank',1:5), 'Frame',
        .1,'s',.1,'s')
```

[Middle]

```
sn = mirzerocross(mirsegment(aTest, 'MFCC', 'Rank',1:5), 'Frame')
```

[Bottom]

```
sb = mirbrightness(mirsegment(aTest, 'MFCC', 'Rank',1:5),
        'Frame', .1,'s',.1,'s')
```

Figure 17

```
aTest = miraudio('50thSetVocalise.aif', 'Extract', 200, 260)
sp  = mirpitch(mirsegment(aTest, 'MFCC',
        'Rank',1:5),'Min',100,'Total',5,'Frame', .1,'s',.1,'s')
```

Figure 18

```
a = miraudio('50thSetVocalise.aif')
segf = mirsegment(a,[75 229.1])
```

Figure 19 [Top and Bottom]

```
a1 = miraudio('50thSetVocalise.aif', 'Extract', 0, 75)
[s p g] = mirsegment(a1,'Contrast',.05, 'KernelSize',70)
```

Figure 20 [Top]

```
a1 = miraudio('50thSetVocalise.aif', 'Extract', 0, 75)
[s p g] = mirsegment(a1,'Contrast',.05, 'KernelSize',70)
c = mirchromagram(a1, 'Frame')
```

[Bottom]

```
cs = mirchromagram(s)
```

Figure 22 [Top Left and Top Right]

```
a1L = miraudio('50thSetVocalise.L.aif', 'Extract', 0, 75)
[sL pL] = mirsegment(a1L,'Contrast',.05, 'KernelSize',70)
```

[Bottom Left and Bottom Right]

```
a1R = miraudio('50thSetVocalise.R.aif', 'Extract', 0, 75)
[sR pR] = mirsegment(a1R,'Contrast',.05, 'KernelSize',70)
```

Figure 23 [Top]

```
a1L = miraudio('50thSetVocalise.L.aif', 'Extract', 0, 75)
csL = mirchromagram(sL)
```

[Bottom]

```
a1R = miraudio('50thSetVocalise.R.aif', 'Extract', 0, 75)
csR = mirchromagram(sR)
```

Figure 23 [Top]

```
a1L = miraudio('50thSetVocalise.L.aif', 'Extract', 0, 75)
```

```
    csL = mirchromagram(sL)
```

[Bottom]

```
    a1R = miraudio('50thSetVocalise.R.aif', 'Extract', 0, 75)
    csR = mirchromagram(sR)
```

Figure 24

```
    a1 = miraudio('50thSetVocalise.aif', 'Extract', 0, 75)
    layerseg = mirsegment(a1,[{0 21},{17.7 42},{38 54.4},{45
        70.3},{70.3 75}])
```

Figure 25

```
    a2 = miraudio('50thSetVocalise.aif', 'Extract', 75, 229.1)
    n = mirnovelty(mirsegment(a2, [93.47 216.2]), 'Frame',
        .1,'s',.1,'s')
```

Figure 26

```
    a2 = miraudio('50thSetVocalise.aif', 'Extract', 75, 229.1)
    c = mirchromagram(a2, 'Frame')
```

Figure 27

```
    a2 = miraudio('50thSetVocalise.aif', 'Extract', 75, 229.1)
    c2 = mirchromagram(a2, 'Frame', 'Wrap', 0)
```

Figure 28 [Left]

```
    a2 = miraudio('50thSetVocalise.aif', 'Extract', 75, 229.1)
    sc = mircentroid(mirsegment(a2, [93.47 216.2]), 'Frame',
        .1,'s',.1,'s')
```

[Right]

```
    ssp = mirspread(mirsegment(a2, [93.47 216.2]), 'Frame',
        .1,'s',.1,'s')
```

Figure 29

```
    fseg = mirsegment(a2, [93.47 216.2])
```

Figure 30

```
    a2L = miraudio('50thSetVocalise.L.aif', 'Extract', 75, 229.1)
    [sL pL gL] = mirsegment(a2L,'Contrast',.10, 'KernelSize',50)
```

Figure 31

```
a2L = miraudio('50thSetVocalise.L.aif', 'Extract', 75, 229.1)
sd = mirroughness(sL, 'Frame', .1,'s',.1,'s')
```

Figure 32

```
a2R = miraudio('50thSetVocalise.R.aif', 'Extract', 75, 229.1)
[sR pR gR] = mirsegment(a2R,'Contrast',.1, 'KernelSize',50)
```

Figure 33 [Top]

```
a2R = miraudio('50thSetVocalise.R.aif', 'Extract', 75, 229.1)
dis = mirroughness(sR, 'Frame', .1,'s',.1,'s')
```

[Bottom]

```
sp = mirspread(sR, 'Frame', .1,'s',.1,'s')
```

Figure 34

```
a2R = miraudio('50thSetVocalise.R.aif', 'Extract', 75, 229.1)
a2RFB = mirfilterbank(a2R, 'Manual',[3000 Inf])
a2RFB2 = miraudio(a2RFB, 'Normal')
c = mirchromagram(a2RFB2, 'Frame', 'Wrap', 0, 'Min',3000,
    'Normal', 1)
```

Figure 35 [Top]

```
a2L = miraudio('50thSetVocalise.L.aif', 'Extract', 75, 229.1)
[sL pL gL] = mirsegment(a2L,'Contrast',.1, 'KernelSize',50)
c = mirchromagram(sL, 'Wrap', 0, 'Frame', .1,'s',.1,'s', 'Max',
    440)
```

[Bottom]

```
a2R = miraudio('50thSetVocalise.R.aif', 'Extract', 75, 229.1)
[sR pR gR] = mirsegment(a2R,'Contrast',.1, 'KernelSize',50)
c = mirchromagram(sR, 'Wrap', 0, 'Frame', .1,'s',.1,'s', 'Max',
    440)
```

Figure 36

```
a2L = miraudio('50thSetVocalise.L.aif', 'Extract', 75, 229.1)
layseg = mirsegment( a2L, [92.8 97.8 102.8 107.92 112.92 118.12
```

```
            122.97 128.17 133.27 138.3 143.37 148.5 151])
    b = mirbrightness(layseg, 'Frame', .1,'s',.1,'s')
```

Figure 37

```
    a2L = miraudio('50thSetVocalise.L.aif', 'Extract', 75, 229.1)
    a2LFB = mirfilterbank(a2L, 'Manual',[700 Inf])
    [sL pL gL] = mirsegment(a2LFB,'Contrast',.1, 'KernelSize',5)
    layseg = mirsegment( a2L, [100.5 102.3 106.67 110.6 114.7 119.1
            123 127.1 131.3 135.4 139.4 143.4 147.67 151.9 156 160.17 164.27 168.4
            172.7 176.9 181.27 185.5 189.8 193.8 197.9 202.3 206.2 210.37 214.57
            218.8 222.9])
    sp = mirspread(layseg, 'Frame', .1,'s',.1,'s')
```

Figure 38

See code for Figure 34

Figure 39[Top]

```
    a2L = miraudio('50thSetVocalise.L.aif', 'Extract', 215, 229.1)
    a2LFB = mirfilterbank(a2L, 'Manual',[-Inf 800])
    [sL pL gL] = mirsegment(a2LFB,[216.2 218.75 223.1 223.75 226.25])
    dis = mirroughness(sL, 'Frame')
```

[Bottom]

```
    c = mirchromagram(sL, 'Frame','Wrap', 0)
```

Figure 40 [Top Left]

```
    a2L = miraudio('50thSetVocalise.L.aif', 'Extract', 200, 270)
    a2LFB = mirfilterbank(a2L, 'Manual',[1000 Inf])
    [sL pL gL] = mirsegment(a2LFB,[216.2 248.75])
    sp = mirspread(sL, 'Frame', .1,'s',.1,'s')
```

[Top Right]

```
    dis = mirroughness(sL, 'Frame', .1,'s',.1,'s')
```

[Bottom Left]

```
    sc = mircentroid(sL, 'Frame', .1,'s',.1,'s')
```

[Bottom Right]

```
c = mirchromagram(sL, 'Frame','Wrap', 0)
```

Figure 41 [Left]

```
a2R = miraudio('50thSetVocalise.R.aif', 'Extract', 205, 240)
a2RFB = mirfilterbank(a2R, 'Manual',[1000 Inf])
[sR pR gR] = mirsegment(a2RFB,[217.6 229.1])
sc = mircentroid(sR, 'Frame', .1,'s',.1,'s')
```

[Right]

```
b = mirbrightness(sR, 'Frame', .1,'s',.1,'s')
```

Figure 42

```
a2R = miraudio('50thSetVocalise.R.aif', 'Extract', 75, 229.1)
a2RFB = mirfilterbank(a2R, 'Manual',[-Inf 1000])
matr = mirsimatrix(a2RFB)
```

Figure 43 [Left]

```
a2R = miraudio('50thSetVocalise.R.aif', 'Extract', 216.2, 229.1)
a2RFB = mirfilterbank(a2R, 'Manual',[-Inf 600])
[sR pR gR] = mirsegment(a2RFB,'Contrast',.1, 'KernelSize',25)
```

[Right]

```
c = mirchromagram(sR, 'Frame','Wrap', 0, 'Normal')
```

Figure 44

```
a3 = miraudio('50thSetVocalise.aif', 'Extract', 229.1, 343.0)
[s p g] = mirsegment(a3,'Contrast',.20, 'KernelSize',25)
```

Figure 45[Top]

```
a3 = miraudio('50thSetVocalise.aif', 'Extract', 229.1, 343.0)
[s p g] = mirsegment(a3,[246.6 255.8 310.4])
b = mirbrightness(s, 'Frame', .1,'s',.1,'s')
```

[Middle]

```
dis = mirroughness(s, 'Frame', .1,'s',.1,'s')
```

[Bottom]

```
sc = mircentroid(s, 'Frame', .1,'s',.1,'s')
```

Figure 46

```
a3L = miraudio('50thSetVocalise.L.aif', 'Extract', 210, 260)
a3LFB = mirfilterbank(a3L, 'Manual',[-Inf 1000])
[s p g] = mirsegment(a3LFB,[229.1 246.6])
dis = mirroughness(s, 'Frame', .1,'s',.1,'s')
```

Figure 47 [Top]

```
a3R = miraudio('50thSetVocalise.R.aif', 'Extract', 210, 260)
[s p g] = mirsegment(a3R,[229.1 246.6])
dis = mirroughness(s, 'Frame', .1,'s',.1,'s')
```

[Bottom]

```
FB = mirfilterbank(a3R, 'Manual',[-Inf 1000])
[s p g] = mirsegment(FB,[229.1 246.6])
dis = mirroughness(s, 'Frame', .1,'s',.1,'s')
```

Figure 48 [Right]

```
a3R = miraudio('50thSetVocalise.R.aif', 'Extract', 240, 320)
[s p g] = mirsegment(a3R,[255.8 310.4])
b = mirbrightness(s, 'Frame', .1,'s',.1,'s')
dis = mirroughness(s, 'Frame', .1,'s',.1,'s')
sc = mircentroid(s, 'Frame', .1,'s',.1,'s')
```

[Left]

```
a3L = miraudio('50thSetVocalise.L.aif', 'Extract', 240, 320)
sp = mirspread(s, 'Frame', .1,'s',.1,'s')
b = mirbrightness(s, 'Frame', .1,'s',.1,'s')
dis = mirroughness(s, 'Frame', .1,'s',.1,'s')
sc = mircentroid(s, 'Frame', .1,'s',.1,'s')
```

Figure 49 [Top]

```
a3 = miraudio('50thSetVocalise.aif', 'Extract', 240, 343.0)
[s p g] = mirsegment(a3,'Contrast',.20, 'KernelSize',25)
```

[Bottom]

```
FB = mirfilterbank(a3, 'Manual',[800 Inf])
```

```
[s p g] = mirsegment(FB,[285.2])
sc = mircentroid(s, 'Frame', .1,'s',.1,'s')
```

Figure 50 [Top]

```
a = miraudio('50thSetVocalise.aif', 'Extract', 60, 85)
FB = mirfilterbank(a, 'Manual',[-Inf 1000])
[s p g] = mirsegment(FB,[70.3 75 75.5 78.5])
```

[Bottom Left]

```
dis = mirroughness(s, 'Frame', .1,'s',.1,'s')
```

[Bottom Right]

```
sp = mirspread(s, 'Frame', .1,'s',.1,'s')
```

Figure 51 [Top]

```
a2L = miraudio('50thSetVocalise.L.aif', 'Extract', 215, 229.1)
a2LFB = mirfilterbank(a2L, 'Manual',[-Inf 800])
[sL pL gL] = mirsegment(a2LFB,[216.2 218.75 223.1 223.75 226.25])
dis = mirroughness(sL, 'Frame')
```

[Bottom]

```
c = mirchromagram(sL, 'Frame','Wrap', 0)
```

Figure 52 [Top]

```
a = miraudio('50thSetVocalise.aif', 'Extract', 230, 270)
FB = mirfilterbank(a, 'Manual',[-Inf 1000])
[s p g] = mirsegment(FB,[246.8 252.1 258.8])
```

[Bottom Left]

```
dis = mirroughness(s, 'Frame', .1,'s',.1,'s')
```

[Bottom Right]

```
sp = mirspread(s, 'Frame', .1,'s',.1,'s')
```